

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

DETEKCE SÍŇOVÝCH EXTRASYSTOL V EKG

PREMATURE ATRIAL CONTRACTION DETECTION IN ECG

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Jana Mistrová

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Petra Novotná

BRNO 2021

Bakalářská práce

bakalářský studijní program **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

Studentka: Jana Mistrová

ID: 211663

Ročník: 3

Akademický rok: 2020/21

NÁZEV TÉMATU:

Detekce síňových extrasystol v EKG

POKYNY PRO VYPRACOVÁNÍ:

1) Nastudujte a popište fyziologii síňových extrasystol v EKG. 2) Nastudujte a popište současné metody detekce síňových extrasystol. Zaměřte se na metody strojového učení. 3) Seznamte se s poskytnutými daty z databáze na ÚBMI a proveďte jejich analýzu. Na základě této analýzy a vybrané detekční metody proveďte potřebné předzpracování dat (filtrace, odvození příznaků, vyhodnocení příznaků, redukce dimenzionality,...). 4) Vybranou detekční metodu podrobně popište a navrhnete způsob její realizace. 5) Vybranou metodu realizujte v prostředí MATLAB. 6) Metodu detekce extrasystol optimalizujte, otestujte a porovnejte a vyhodnoťte její přínosy a nedostatky. 7) Dosažené výsledky vyhodnoťte a podrobně diskutujte.

DOPORUČENÁ LITERATURA:

[1] LUZ, Eduardo José da S., William Robson SCHWARTZ, Guillermo CÁMARA-CHÁVEZ a David MENOTTI. ECG-based heartbeat classification for arrhythmia detection: A survey. Computer Methods and Programs in Biomedicine. 2016, (127), 144-164. DOI: 10.1016/j.cmpb.2015.12.008. ISSN 01692607.

[2] GOLDBERGER, Ary L., Zachary D. GOLDBERGER a Alexei SHVILKIN. Goldberger's clinical electrocardiography: a simplified approach. 9th edition. Philadelphia: Elsevier, 2018. ISBN 978-0-323-40169-2.

Termín zadání: 8.2.2021

Termín odevzdání: 28.5.2021

Vedoucí práce: Ing. Petra Novotná

doc. Ing. Jana Kolářová, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá problematikou detekce síňových extrasystol z EKG signálu. V první části práce je popsán elektrokardiogram, převodní systém srdeční a extrasystoly. Extrasystoly jsou předčasné srdeční stahy, které vznikají v ektopických ložiskách. Další části jsou věnovány předzpracování EKG signálu, výběru příznaků, jejich redukci a klasifikačním metodám. V poslední části je realizovaná metoda detekce a výsledky klasifikace.

KLÍČOVÁ SLOVA

elektrokardiogram, EKG, síňová extrasystola, detekce síňových extrasystol

ABSTRACT

This thesis discusses detection of premature atrial contraction from ECG. In the first part, thesis describes electrocardiogram, cardiac conduction system and extrasystoles. Extrasystoles are premature contraction caused by ectopic heartbeats. Next part is devoted to signal preprocessing, the feature description methods, reduction of feature vector and methods of classification. Realized method and results of classifier are discussed in the last part.

KEYWORDS

electrocardiogram, ECG, premature atrial contraction, detection of premature atrial contraction

MISTROVÁ, Jana. *Detekce síňových extrasystol v EKG*. Brno, 2020, 49 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce: Ing. Petra Novotná

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Detekce síňových extrasystol v EKG“ jsem vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autorka uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autorky

Obsah

Úvod	9
1 Činnost srdce	10
1.1 Převodní systém srdeční	10
1.2 Elektrokardiografie	11
1.2.1 Křivka EKG	12
1.3 Extrasystoly	12
1.3.1 Komorové extrasystoly	13
1.3.2 Junkční extrasystoly	14
1.3.3 Síňové extrasystoly	14
2 Předzpracování EKG signálu	17
2.1 Rušení a jeho filtrace	17
2.2 Segmentace	18
3 Selekcce příznaků	20
3.1 RR intervaly	20
3.2 Intervalové příznaky	20
3.3 Morfologické příznaky	20
3.4 Diskrétní vlnková transformace (DWT)	21
4 Redukce příznaků	23
4.1 Analýza hlavních komponent (PCA)	23
4.2 Analýza nezávislých komponent (ICA)	24
5 Klasifikátory	25
5.1 Podpůrné vektory	25
5.2 Neuronové sítě	25
5.3 Lineární diskriminant	27
5.4 K-nejbližší soused	27
6 Praktická část práce	28
6.1 Vstupní data	28
6.1.1 Vybraná množina dat a její rozdělení	28
6.1.2 Filtrace dat	28
6.1.3 Detekované R kmity	29
6.1.4 Segmentace dat	30
6.1.5 Výpočet příznaků	32

7	Výsledky klasifikace metodou k-nejbližší soused	36
7.1	Kombinace různých příznaků	38
7.2	Metoda PCA	40
8	Diskuze	42
	Závěr	44
	Literatura	46
	Seznam symbolů, veličin a zkratk	49

Seznam obrázků

1.1	Převodní systém srdeční, převzato z: [7]	11
1.2	Einthovenův trojúhelník (vlevo) a umístění Wilsonových hrudních svodů (vpravo), převzato z: [4]	11
1.3	Křivka EKG, převzato z: [13]	12
1.4	Třetím srdečním stahem je komorová extrasystola se širokým QRS komplexem, převzato z: [13]	14
1.5	Signál s junkčními extrasystolami, převzato z: [13]	14
1.6	Po druhém a desátém sinusovém rytmu následuje síňová extrasystola s neúplnou kompenzační pauzou	15
1.7	Sinusový rytmus s blokádou levého Tawarova raménka, převzato z: [13]	16
1.8	Blokovaná síňová extrasystola, převzato z: [13]	16
2.1	Signál zatížený driftem	17
2.2	Signál zatížený brumem	18
3.1	Dvě metody vzorkování pro výběr příznaků. První metoda (vlevo) se vzorkovacími okny QRS onset - QRS offset a QRS offset - T-wave offset. Druhá metoda (vpravo) se vzorkovacími okny -50 ms až +100 ms a +150 ms až 500 ms od kmitu R. Převzato z: [14]	21
5.1	Model perceptronu, převzato z: [17]	26
5.2	Vícevrstvá perceptronová síť se dvěmi skrytými vrstvami, převzato z:[16]	26
6.1	Signál zatížený driftem (vlevo), signál zatížený brumem (vpravo) . . .	29
6.2	Signál s odfiltrovaným brumem a driftem, bez vyhlazení (vlevo), signál s filtrem na vyhlazení (vpravo)	29
6.3	Příklad chybné detekce R kmitů, první R kmit nebyl označen	30
6.4	Špatně detekovaný R kmit (vlevo) a posunutá hodnota R kmitu na maximum (vpravo)	30
6.5	Správně detekované Q a S body u sinusového EKG rytmu	31
6.6	Chybná detekce Q a R bodu u komorové extrasystoly a správně na-detekované body QRS komplexu u síňové extrasystoly	32
6.7	Odhad začátku a konce QRS komplexu u sinusového EKG označené křížkem	34
6.8	Označené začátky a konce QRS komplexu u signálu se síňovou extrasystolou.	34
7.1	Přesnost detekce v závislosti na počtu zvolených sousedů u klasifikátoru pro výpočet se všemi příznaky	37
7.2	Přesnost klasifikátoru v závislosti na počtu hlavních komponent pro 4. kombinaci příznaků	41

Seznam tabulek

6.1	Vybrané příznaky	33
7.1	Výsledky klasifikace pro příznakový vektor s počtem sousedů pro lichá čísla od 1 do 17	37
7.2	Vybrané kombinace příznaků	38
7.3	Výsledky klasifikace pro různé varianty nakombinovaných příznaků s jedním sousedem	38
7.4	Výsledky klasifikace s jedním sousedem po aplikaci PCA pro varianty příznaků	40
8.1	Výsledky klasifikace pomocí k-NN a s PCA pro všechny příznaky v horní části tabulky a pro 5. kombinaci příznaků v dolní části tabulky	42
8.2	Nejlepší dosažené výsledky klasifikace pomocí k-NN a s PCA pro 4. kombinaci příznaků	43
8.3	Výsledky klasifikace pomocí k-NN a s PCA pro 3. kombinaci příznaků v horní části tabulky a pro 7. kombinaci příznaků v dolní části tabulky	43

Úvod

Tato bakalářská práce se zabývá klasifikací síňových extrasystol v elektrokardiografických záznamech (EKG).

Výskyt síňových extrasystol je v EKG záznamech velice častý. Vyskytují se i u zdravých jedinců a neznamenaají pro ně žádné nebezpečí. Existují ovšem i případy, kdy přítomnost síňových extrasystol může znamenat závažný problém, protože souvisí s vážnou srdeční chorobou a nebo jejich častý výskyt může vést k jiným poruchám srdečního rytmu. Z tohoto důvodu je důležité správně a včas detekovat jejich přítomnost.

Součástí práce je podrobný popis fyziologie extrasystol, která je nutná pro následnou klasifikaci. V dalších kapitolách práce jsou popsány postupy při výběru příznaků, jejich redukce a následně klasifikace pomocí klasifikačních metod, které jsou dnes hojně využívány.

Další částí bakalářské práce je navržený algoritmus, který detekuje síňové extrasystoly na základě vypočtených příznaků. Je využita metoda PCA pro redukci příznakového vektoru s následnou klasifikací pomocí metody k-nejbližšího souseda.

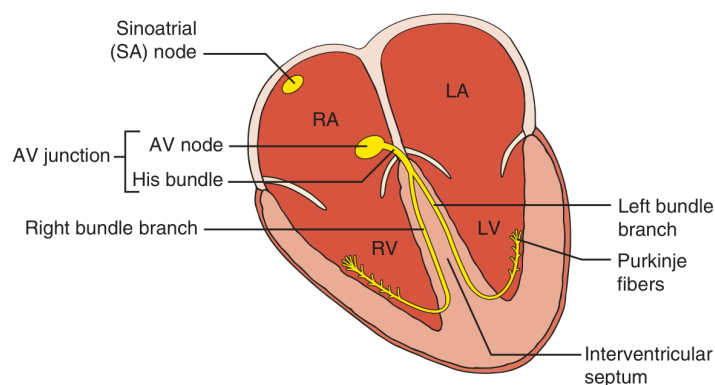
1 Činnost srdce

Srdce je dutý svalový orgán, který je uložen v hrudníku. Srdce je tvořeno čtyřmi dutinami: levou síní, pravou síní, levou komorou a pravou komorou. Mezi levou síní a levou komorou se nachází dvojcípá chlopeč, která zprostředkovává tok krve ze síně do komory. Stejně tak u pravé síně a pravé komory, které jsou od sebe odděleny trojcípou chlopní. Při otevření poloměsíčité chlopně je krev z levé srdeční komory vypuzena tepnami do tělního oběhu. Naopak žíly vedou krev zpět do srdce a ta je pravou srdeční komorou vypuzena do plicního oběhu a následně přivedena do levé síně [1].

Elektrickým podrážděním srdce dochází k činnosti srdečních komor a k pravidelnému střídání fáze systoly a diastoly. Diastola nastává fází izovolumické relaxace, kdy tlak v síních stoupá a to způsobí otevření cípatých chlopní. Další fází diastoly je fáze plnění, kdy krev teče ze síní do komor. Ke konci diastoly dochází ke kontrakci síní. Cípaté chlopně se uzavřou, když je nitrokomorový tlak vyšší než tlak v síních. Jakmile nitrokomorový tlak v levé komoře překročí tlak v aortě, otevřou se poloměsíčité chlopně a začíná vypuzovací fáze. Během této fáze je krev vypuzena ze srdečních komor a tlak v levé komoře klesá pod hodnotu tlaku v aortě a tím dochází k uzavření poloměsíčitých chlopní. Poté nastává opět plnicí fáze a srdeční cyklus se opakuje [1].

1.1 Převodní systém srdeční

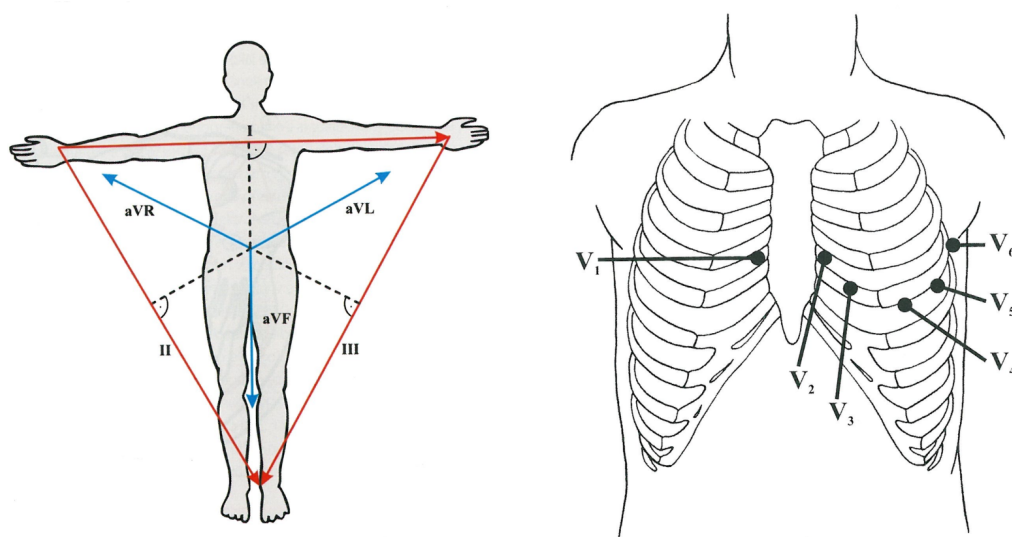
Převodní systém srdeční je tvořen svalovými buňkami myokardu a některé z nich dokáží tvořit a vést vzruchy. Vzruch fyziologicky vzniká v sinoatriálním (SA) uzlu, který je umístěn ve stěně pravé síně. Jedná se o srdeční pacemaker, protože vytváří pravidelné vzruchy o frekvenci srdečních stahů 60-100/min. Schopnost tvořit vzruchy mají i ostatní oddíly převodního systému, ale s odlišnou frekvencí. Proto se za normálních podmínek vzruchy tvoří v SA uzlu, který má nejvyšší frekvenci tvorby. Tento podnět k podráždění se z SA uzlu šíří do atrioventrikulárního (AV) uzlu, který je umístěn v dolní části pravé síně. Vedení vzruchu je v AV uzlu zpožděno z důvodu dokončení kontrakce síní. Dále se šíří Hisovým svazkem, který se rozděluje na levé a pravé Tawarovo raménko, až k Purkyňovým vláknům, která přenáší podráždění na levou a pravou komoru [1]. Převodní systém srdeční je zobrazen na obrázku 1.1.



Obr. 1.1: Převodní systém srdeční, převzato z: [7]

1.2 Elektrokardiografie

Při elektrokardiografii se měří elektrická aktivita srdce a vyhodnocuje se pomocí elektrokardiogramu (EKG), který zobrazuje potenciálové rozdíly vznikající při podráždění srdce. Nejčastěji se k měření používá 12svodové EKG za použití hrudních a končetinových svodů. Einthovenovy bipolární končetinové svody I, II a III jsou umístěny na levé horní a dolní končetině a na zápěstí pravé horní končetiny. Na pravé dolní končetině se nachází zemní elektroda. Tyto svody dohromady tvoří Einthovenův trojúhelník, který lze vidět na obrázku 1.2. EKG je dále měřeno pomocí unipolárních Goldbergerových svodů aVL, aVR a aVF, kde je aktivní elektroda umístěna vždy na jedné končetině. Mezi další unipolární svody, které jsou součástí 12svodového měření, patří hrudní Wilsonovy svody V1-V6 [1, 20].



Obr. 1.2: Einthovenův trojúhelník (vlevo) a umístění Wilsonových hrudních svodů (vpravo), převzato z: [4]

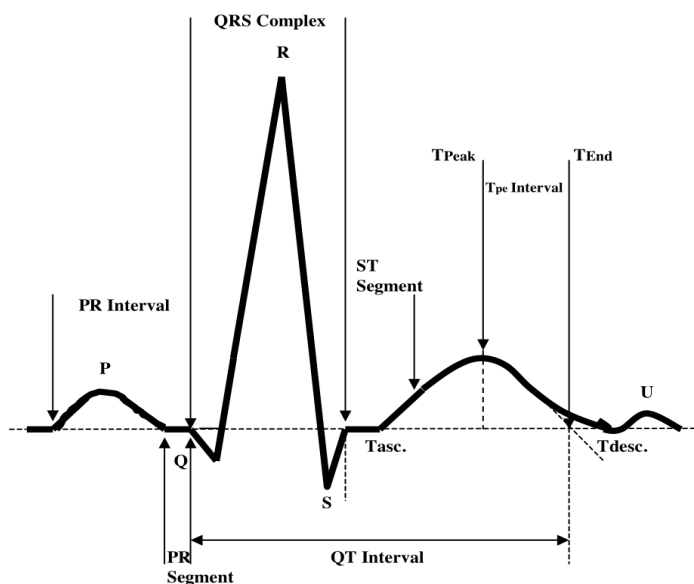
1.2.1 Křivka EKG

Křivka EKG zobrazuje srdeční aktivitu, kde pozorujeme jednotlivé srdeční cykly. Na této křivce rozpoznáváme kmity, vlny, úseky a intervaly, které lze vidět na obrázku 1.3 [1, 8].

Vlna P vyjadřuje depolarizaci síní, která je způsobena vedením podráždění z SA uzlu na síně. PQ interval znázorňuje převod vzruchu ze síní na komory AV uzlem a trvá 0,12-0,2 s. K repolarizaci síní dochází v době průběhu QRS komplexu [1, 8].

Následuje QRS komplex, který je tvořen negativním kmitem Q, pozitivním kmitem R a dalším negativním kmitem S. Časově odpovídá 0,08 s. Komplex QRS zobrazuje vedení vzruchu přes AV uzel, Hisův svazek až na Purkyňova vlákna, což zapříčiní depolarizaci komor [1, 8].

Vlna T je projevem repolarizace komor. ST úsek vyjadřuje kompletně aktivované komory. Interval QT odpovídá celkovému trvání depolarizace a repolarizace komor a jeho doba trvání je okolo 0,35 s [1, 8].



Obr. 1.3: Křivka EKG, převzato z: [13]

1.3 Extrasystoly

Extrasystola znamená předčasný srdeční stah, který přichází před očekávanou srdeční kontrakcí. Jedná se tedy o stah mimo pravidelný srdeční rytmus. Tento impulz pro předčasnou srdeční kontrakci vzniká v ložisku v srdci mimo SA uzel a dochází k předčasné depolarizaci myokardu. Předčasně vytvořený vzruch tak předbíhá

vzruch z SA uzlu. Extrasystoly na základě místa jejich vzniku rozdělujeme na supra-ventrikulární (síňové, junkční) a komorové [21].

Vznik extrasystol nemá většinou jasnou příčinu a běžně se vyskytují u pacientů s pravidelným srdečním rytmem jakéhokoli věku. Takový pacient občasnou přítomnost předčasných stahů nepocituje, ovšem při jejich častějším výskytu může pacient zaznamenat bolest na hrudi, pocit vynechání srdce nebo bušení srdce. Extrasystoly se v srdečním rytmu nacházejí častěji s přibývajícím věkem pacienta. K vyššímu výskytu extrasystol může přispět nejen nadměrná konzumace kofeinu, alkoholu nebo nikotinu, ale také stres, hypertenze, porucha vnitřního prostředí nebo hypertyreóza [2, 21].

Ve většině případů nejsou příznakem srdečního onemocnění, ale existují případy, kdy předčasné srdeční stahy souvisí s vážným srdečním onemocněním, jako například ischemickou chorobou srdeční nebo kardiomyopatií. Častý výskyt síňových extrasystol může vést například k fibrilaci síní, flutteru síní a nebo může způsobovat další supraventrikulární tachyarytmie, jako jsou atrioventrikulární reentry tachykardie nebo atrioventrikulární nodální reentry tachykardie [2, 21].

Extrasystoly se mohou vyskytovat v páru se sinusovým rytmem. Na záznamu je pak každý sinusový rytmus následován extrasystolou. To se nazývá bigeminie. Každý třetí předčasný stah je označován jako trigeminie. Kvadrigeminie vyjadřuje extrasystolu po třech sinusových srdečních stazích. Dvě extrasystoly nacházející se za sebou jsou označovány jako kuplet a tři extrasystoly za sebou jako triplet [2].

1.3.1 Komorové extrasystoly

Vzruch pro vznik komorových extrasystol vzniká v myokardu komor a dochází k tomu, že oblast komory, kde vznikl podnět k podráždění, je depolarizována dříve než zbytek komory a druhá komora se depolarizuje až následně. To způsobuje změny na křivce EKG (obrázek 1.4). Vlna P se na záznamu neobjevuje. QRS komplex je oproti tomu fyziologickému abnormálně široký a spolu s vlnou T se obvykle vyskytuje v negativních hodnotách, tedy v opačném směru [4, 7].

Na záznamu EKG se ve většině případů objevuje úplná kompenzační pauza. Vzruch je vytvořen ve svalovině komor a často dochází k tomu, že je veden retrográdně (zpětně) z komor na síně. Toto vedení je ale přerušeno Hisovým svazkem a Purkyňovými vlákny, nebo při střetnutí se vzruchem z SA uzlu. Po komorové extrasystole se totiž stále v SA uzlu tvoří vzruchy v pravidelném rytmu (zjevná vlna P), ale k depolarizaci komor nedochází, protože AV uzel je v refrakterní fázi a nerozšíří impuls do komor. Následuje úplná kompenzační pauza. Její délka je dvojnásobek sinusového RR intervalu, tedy vzdálenost QRS komplexu před a po komorové extrasystole [7].



Obr. 1.4: Třetím srdečním stahem je komorová extrasystola se širokým QRS komplexem, převzato z: [13]

1.3.2 Junkční extrasystoly

Junkční extrasystoly jsou předčasné srdeční kontrakce, které vznikají ve svalovině síní stejně, jako síňové extrasystoly. Junkční extrasystoly vznikají konkrétně v AV uzlu, odtud je vzruch dále šířen do myokardu komor, takže na EKG křivce lze pozorovat normálně štíhlý QRS komplex. Vzhledem k tomu, že depolarizace síní neprobíhá standardně, P vlna se na EKG křivce vůbec neobjevuje, je skryta v QRS komplexu, a nebo se výjimečně vyskytuje za ním z důvodu retrográdního vedení [4, 20]. Příklad signálu s junkčními extrasystolami je uveden na obrázku 1.5.



Obr. 1.5: Signál s junkčními extrasystolami, převzato z: [13]

1.3.3 Síňové extrasystoly

V ektopickém ložisku v jedné ze síní a nebo v interatriálním septu (mimo SA uzel) se vytvoří akční potenciál dříve než je vytvořen v SA uzlu a tím dochází k síňovým extrasystolám. Dále je vzruch šířen přes AV uzel, Hisův svazek a Purkyňova vlákna do komor, které jsou normálně depolarizovány. Na křivce EKG je tedy fyziologický QRS komplex. Na rozdíl od komorových extrasystol, které mají široký QRS komplex [7].

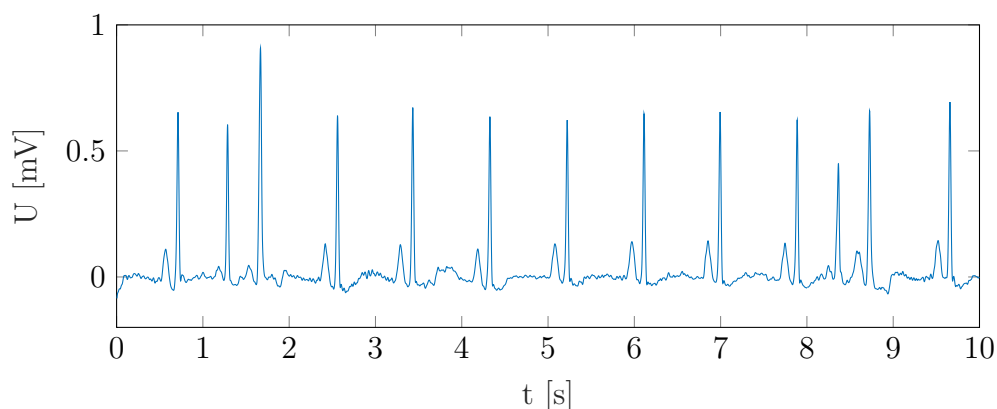
V případě síňových extrasystol jsou srdeční síně předčasně depolarizovány, a to se projeví především na vlně P. Vlna P bude mít odlišný tvar oproti vlně P sinusového rytmu. Tvar vlny P závisí na tom, kde přesně se ektopické ložisko vznikajícího vzruchu nachází. Pokud je v blízkosti SA uzlu, bude mít P vlna podobný tvar jako

sinusová P vlna. V případě, že k vytvoření vzruchu došlo blízko AV uzlu, síně budou depolarizovány v opačném směru a vlna P bude negativní. Předčasná depolarizace síní se na EKG záznamu může projevit i změnou délky PR intervalu. Pokud by byl PR interval delší než obvykle, znamená to, že předčasný vzruch dorazil na AV uzel dříve, než byl kompletně depolarizován a z části se ještě nachází v refrakterní fázi. V případě blokových extrasystol může být P vlna skryta ve vlně T předcházejícího cyklu [7, 15].

Na základě umístění ložisek, kde se tvoří předčasné impulzy, můžeme síňové extrasystoly rozdělit na unifokální a multifokální. V případě, že existuje jedno ek-topické ložisko pro tvorbu vzruchů, pak jsou extrasystoly totožné a jedná se o extrasystoly unifokální. Pokud jsou tato ložiska dvě nebo jich je více, pak vznikají extrasystoly s navzájem odlišnou P vlnou a jsou označovány jako multifokální [2].

Neúplná kompenzační pauza

Pro síňové extrasystoly je typická neúplná kompenzační pauza následující po síňové extrasystole, znázorněna na obrázku 1.6. Tato pauza nastává, pokud impuls pro předčasný síňový stah vybijí SA uzel, než se stihne vytvořit sinusový vzruch. Neúplná kompenzační pauza je vzdálenost QRS komplexů před a za extrasystolou a její délka je menší než dvojnásobek sinusového RR intervalu. Doba, o kterou je tato vzdálenost QRS komplexů kratší než normálně, je dána tím, o kolik dříve předčasný impuls vybijí vzruch, který se pravidelně tvoří v SA uzlu [7, 15].



Obr. 1.6: Po druhém a desátém sinusovém rytmu následuje síňová extrasystola s neúplnou kompenzační pauzou

Síňová extrasystola s aberantním vedením po komorách

V případě aberantního komorového vedení dochází u extrasystol k širokému QRS komplexu s abnormálním tvarem (obrázek 1.7). Je to způsobeno bloádou levého nebo

pravého Tawarova raménka. Když předčasný impuls zastihne jedno z Tawarových ramének v refrakterní fázi, dochází k jeho zablokování. Na EKG je pak viditelná P vlna předčasného stahu a za ní široký QRS komplex. Síňové extrasystoly s aberantním QRS komplexem mohou být omylem zaměněny s komorovými extrasystolami, které se také vyznačují širokým QRS komplexem [15].



Obr. 1.7: Sinusový rytmus s blokádou levého Tawarova raménka, převzato z: [13]

Blokované síňové extrasystoly

Blokované síňové extrasystoly vznikají, pokud byl vytvořen velmi předčasný vzruch těsně po předcházejícím vzruchu, tím pádem byl AV uzel uveden do refrakterní fáze. Procházející vzruch nemůže být tedy převeden na komory a nedochází k jejich depolarizaci. Na EKG bude abnormální vlna P, a nebo tato vlna může být součástí vlny T předcházejícího rytmu, pak by tato vlna T vypadala odlišně oproti fyziologické. Vlna P není následována komplexem QRS ani T vlnou [7, 15]. Blokovaná síňová extrasystola je zobrazena na obrázku 1.8



Obr. 1.8: Blokovaná síňová extrasystola, převzato z: [13]

2 Předzpracování EKG signálu

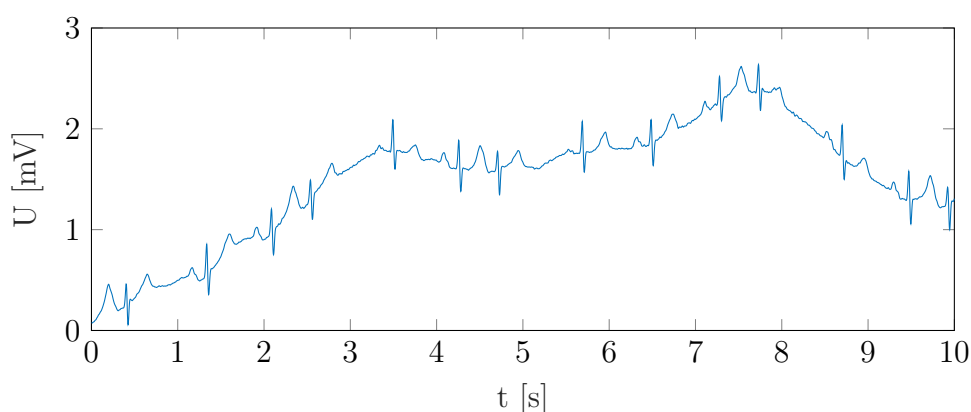
2.1 Rušení a jeho filtrace

Při snímání EKG signálu dochází ke vzniku různých typů rušení, které poškozují signál a není možné provést jeho diagnostiku nebo zpracování. Mezi taková rušení patří kolísání nulové izolinie (drift) a síťové rušení (brum), které je nutné potlačit i při detekci síňových extrasystol [9].

Drift a jeho filtrace

Spektrum nízkofrekvenčního driftu se nejčastěji pohybuje kolem frekvence 0,5 Hz, obecně se uvádí až do 2 Hz. Kolísání nulové izolinie, které znehodnocuje signál, je způsobeno tím, že spektrum rušení se navzájem překrývá s užitečnou složkou signálu. Dochází tedy k potlačování informací, které jsou součástí EKG signálu, proto je nutné drift potlačit, aby bylo umožněno následné zpracování signálu. Kolísání nulové izolinie může být způsobeno pohybem pacienta, elektrochemickými procesy mezi elektrodou a pokožkou pacienta a dále jeho dýcháním [9].

Většina prací využívá k filtraci driftu mediánové filtry. Autoři prací [5, 12, 14] nejprve signály filtrovali mediánovým filtrem s délkou okna 200 ms k odstranění P vln a QRS komplexů a následně dalším mediánovým filtrem, avšak o délce okna 600 ms. Tento dvakrát filtrovaný signál je poté odečten od původního signálu, aby výsledný signál nebyl zatížen kolísáním nulové izolinie. Signál zatížený driftem lze vidět na obrázku 2.1.

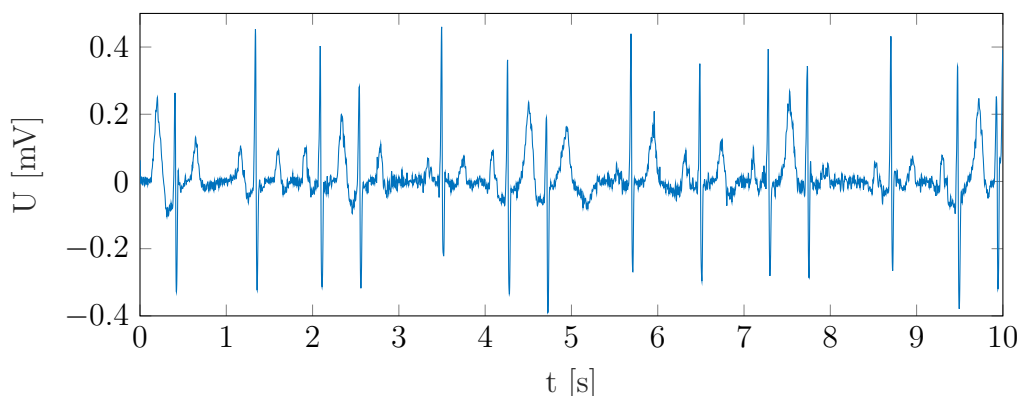


Obr. 2.1: Signál zatížený driftem

Síťové rušení a jeho filtrace

Síťový brum je harmonické periodické rušení, které zatěžuje signál na frekvenci kolem 50 Hz, kdy na této frekvenci dochází k prolínání spektra šumu s užitečnou složkou. Příčinou síťového rušení může být například nevyhovující uzemnění EKG elektrod, pacienta a nebo elektromagnetické rušení z elektrického vedení [9].

Filtrace rušení je nejčastěji realizována pomocí FIR filtrů, tedy filtrů s konečnou impulzní charakteristikou. Pomocí těchto filtrů jde snadno filtrovat například síťové rušení (brum) na 50 Hz, kdy je toto pásmo v okolí 50 Hz vyjmuto. Takovým filtrem může být pásmová zádrž. Dochází i k filtraci nízkých a vysokých frekvenčních složek, které jsou potlačeny filtry horní nebo dolní propusti. K odstranění šumu se mohou dále použít adaptivní filtry nebo vlnková transformace. Dalším využitím adaptivních filtrů je například při segmentaci signálu. V případě vlnkové transformace je to výběr příznaků [13]. Podle studií [12, 14] je dále signál filtrován dolní propustí 12. řádu s mezní frekvencí 35 Hz, který v signálu potlačí síťový brum a vysokofrekvenční rušení. Studie [5] uvádí, že odstranění síťového brumu bylo dosaženo filtrací pásmovou zádrží na frekvenci 60 Hz, v našich podmínkách by se jednalo o frekvenci 50 Hz. Příklad signálu zatíženého brumem je na obrázku 2.2.



Obr. 2.2: Signál zatížený brumem

2.2 Segmentace

Segmentace znamená oddělení jednotlivých úseků signálu, které z hlediska prováděné detekce obsahují užitečné informace. V tomto případě se konkrétně jedná o úsek před a za R kmitem. Pro oddělení této užitečné části je vytvořeno okno obsahující tyto části signálu a následně jsou porovnávány s ostatními okny [9].

Součástí segmentace je také detekce R kmitů, která je nezbytná pro předzpracování EKG signálu. Nízká přesnost detekovaných R kmitů může negativně ovlivnit

výslednou klasifikaci, pokud se chyba šíří do následujících fází zpracování signálu [13]. Tato práce se dále nebude zabývat metodami pro detekci R kmitů, jelikož dostupné signály z databáze již obsahují detekované pozice QRS komplexů.

3 Selekce příznaků

Ze signálu EKG jsou vybírány příznaky, které naznačují přítomnost síňových extrasystol. Vhodný výběr příznaků pro danou problematiku vede k úspěšné klasifikaci arytmií [13].

3.1 RR intervaly

RR intervaly značí vzdálenost dvou R kmitů od sebe. Z této skupiny jsou nejčastěji vybírány čtyři příznaky. Jedná se o pre-RR interval, post-RR interval, průměrný RR interval a lokální průměrný RR interval. Pre-RR interval označuje RR interval mezi daným R kmitem a předchozím, post-RR interval mezi daným R kmitem a následujícím. Průměrný RR interval je průměrem všech RR intervalů ze záznamu a lokální průměrný RR interval je průměr například deseti RR intervalů v okolí daného R kmitu [5, 14].

Ve studii [11] byly použity tyto čtyři příznaky vycházející z RR intervalů. Pre-RR a post-RR stejně jako ve výše uvedených studiích. Dále se jednalo o průměr RR intervalů vypočtených z jednominutového EKG záznamu od daného R kmitu. Následně průměr RR intervalů z dvaceti minutového záznamu.

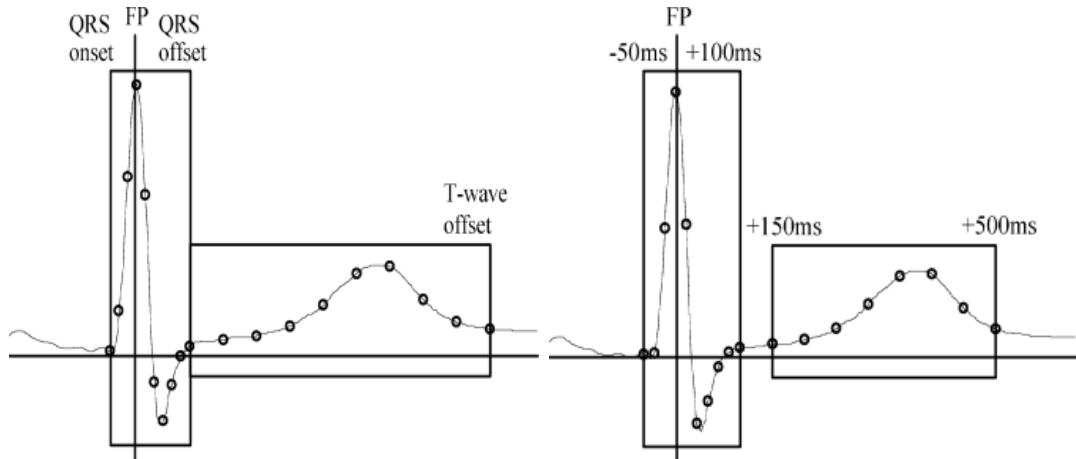
3.2 Intervalové příznaky

Mezi intervalové příznaky řadíme dobu trvání QRS komplexu, dobu trvání T vlny a fakt, zda je přítomná P vlna. Trvání QRS komplexu je vymezeno začátkem a koncem QRS (QRS onset a QRS offset). Podobně trvání T vlny, je určena od konce QRS komplexu až do konce T vlny (QRS offset a T-wave offset [14].

3.3 Morfologické příznaky

Na základě studií [5, 14] byly morfologické příznaky vybírány dvěma metodami pomocí vzorkování. Nutná je předchozí detekce R kmitů. Na vzorkování byla použita dvě vzorkovací okna. U první metody (obrázek 3.1 vlevo) je první vzorkovací okno určeno počátkem a koncem QRS komplexu (QRS onset a QRS offset), který je popsán deseti rovnoměrně rozmístěnými vzorky získaných vzorkováním amplitudy. Druhé vzorkovací okno je popsáno devíti vzorky od konce QRS komplexu do konce T vlny (T-wave offset). Pro odhad mezilehlých hodnot byla použita lineární interpolace. Druhá metoda vzorkování (obrázek 3.1 vpravo) stanovuje první vzorkovací

okno v rozmezí mezi -50 ms od kmitu R a +100 ms od R kmitu a signál je převzorkován frekvencí 60 Hz. Z tohoto okna je následně použito deset vzorků. Druhé okno je +150 ms od R kmitu až do +500 ms od R kmitu a je popsáno osmi vzorky. Na vzorkování je použita nižší vzorkovací frekvence z důvodu nižšího frekvenčního rozsahu než v okně, který zahrnuje QRS komplex.



Obr. 3.1: Dvě metody vzorkování pro výběr příznaků. První metoda (vlevo) se vzorkovacími okny QRS onset - QRS offset a QRS offset - T-wave offset. Druhá metoda (vpravo) se vzorkovacími okny -50 ms až +100 ms a +150 ms až 500 ms od kmitu R. Převzato z: [14]

3.4 Diskrétní vlnková transformace (DWT)

Vlnková transformace je považována za rozšíření Fourierovy transformace, která pracuje pouze ve spektrální oblasti. Výhodou vybírání příznaků pomocí vlnkové transformace je, že získáme informace z frekvenční i časové oblasti. Takovéto práce poté dosahují lepších výsledků a celkově je tato metoda považována za jednu z nejlepších pro výběr příznaků. Nejoblíbenějším typem je diskrétní vlnková transformace [3, 13].

U spojitě vlnkové transformace je mateřská vlnka definována následujícím vztahem 3.1:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi \frac{t - \tau}{s}, \quad (3.1)$$

jedná se o vlnkovou funkci v čase t , ze které se vytvářejí další vlnky změnou měřítka s a časovým posunutím τ .

Mateřská vlnka DWT je popsána níže uvedeným vztahem 3.2:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{s_0^m}} \psi \frac{t - n\tau_0 s_0^m}{s_0^m}, \quad (3.2)$$

do vzorce 3.1 jsou dosazeny diskretizované hodnoty získané logaritmickou diskretizací hodnot τ a s . Jejich diskrétní podoby jsou $s = s_0^m$ a $\tau = n\tau_0 s_0^m$, kde m a n jsou celočíselné konstanty.

Nejčastějším způsobem vzorkování je dyadické, kdy proměnná s_0 nabývá hodnoty 2 a proměnná τ je rovna 1. Po dosazení a úpravě rovnice 3.2 získáme:

$$\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m}t - n). \quad (3.3)$$

Po definování mateřské vlnky může být zápsán vzorec pro DWT 3.4:

$$w_{m,n} = \sum_t f(t) \psi_{m,n}(t), \quad (3.4)$$

kde $w_{m,n}$ reprezentuje vlnkové koeficienty a $f(t)$ je analyzovaný signál [25].

4 Redukce příznaků

Redukce příznaků slouží ke zmenšení velikosti příznakových vektorů nejen z důvodu snížení výpočetní a časové náročnosti při klasifikaci, ale také v případě přítomnosti nerelevantních příznaků pro řešený problém a nebo příznaků na sobě závislých. V případě použití klasifikačních metod, jako jsou podpůrné vektory, k-nejbližší soused nebo neuronové sítě, se s počtem příznaků zvětšuje výpočetní náročnost těchto klasifikátorů, kterou je možné snížit právě zredukováním příznaků. Kdybychom použili nerelevantní příznaky, může dojít ke zmatení klasifikátoru nedůležitými příznaky pro daný problém a tím i ke zhoršení klasifikačních výsledků [13].

4.1 Analýza hlavních komponent (PCA)

Analýza hlavních komponent je statistická metoda, která využívá lineární transformaci k redukci počtu příznaků. Pokud je mezi příznaky vzájemná závislost a data jsou korelovaná, transformuje původní příznaky na příznaky nové nekorelované, které označujeme jako hlavní komponenty. Hlavní komponenty jsou dány velikostí rozptylu, které obsahují většinu informací o variabilitě dat. Jsou vypočítány pomocí vlastních čísel a vektorů. Hlavní komponenty jsou vybírány na základě velikosti rozptylu a vzájemné ortogonalitě. První komponenta vykazuje největší míru proměnlivosti dat. Následuje druhá hlavní komponenta, která má druhý největší rozptyl. A další hlavní komponenty jsou řazeny od maximálního rozptylu po nejmenší [18].

Vzájemná závislost mezi jednotlivými příznaky je určena kovarianční, popř. korelační maticí. Pro výpočet kovariační matice A je nutné pracovat s centrovanými daty, která získáme odečtením střední hodnoty příznaků od každého příznaku daného objektu. Centrování dat je prováděno pomocí vzorce 4.1:

$$y_{i,j} = x_{i,j} - \mu_j, \quad (4.1)$$

kde $y_{i,j}$ je centrovaná hodnota objektu i a příznaku j , x je původní hodnota na pozici i, j a μ je střední hodnota všech příznaků daného objektu. Kovarianční matice A je spočtena pomocí následujícího vzorce 4.2:

$$A = \frac{1}{M-1} Y^T Y, \quad (4.2)$$

kde M je počet objektů, Y je centrovaná matice dat a Y^T je transponovaná matice Y . V případě výpočtu korelační matice provádíme normalizaci všech hodnot kovarianční matice [3, 18].

Dále následuje výpočet vlastních čísel a vlastních vektorů kovarianční matice pomocí spektrálního rozkladu matice A , kdy platí rovnost podle vztahu 4.3:

$$AV = \lambda V, \quad (4.3)$$

kde λ vyjadřuje vlastní číslo a V je matice vlastních vektorů. Pro samotný výpočet vlastních čísel je použit vztah 4.4,

$$\det(A - \lambda I) = 0, \quad (4.4)$$

kde I je jednotková matice. Níže uvedená rovnice 4.5 popisuje dosazení do rovnice 4.4,

$$\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0. \quad (4.5)$$

Po získání hodnot vlastních čísel λ provedeme výpočet vlastních vektorů V podle vztahu 4.6,

$$(A - \lambda I) V = 0, \quad (4.6)$$

následuje obecné dosazení do rovnice 4.6 [18],

$$\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0. \quad (4.7)$$

4.2 Analýza nezávislých komponent (ICA)

Analýza nezávislých komponent je statistická metoda, která se snaží oddělit signálové složky, které jsou statisticky nezávislé. Nezávislé komponenty jsou totiž jakési skryté vlastnosti signálu, které nelze pozorovat během měření. Metoda ICA vyjadřuje lineární kombinace statisticky nezávislých příznaků pomocí vztahu 4.8:

$$x(t) = As(t), \quad (4.8)$$

kde $x(t)$ je původní signál, $s(t)$ představuje výsledné separované složky a A je označení pro transformační matici s transformačními koeficienty. Proměnná $x(t)$ je modelována jako lineární kombinace náhodných proměnných $s(t)$. Vzhledem k tomu, že proměnné $s(t)$ a A nejsou známy a mají být odhadnuté z původních signálů, metoda ICA pracuje s předpokladem, že $s(t)$ jsou vzájemně nezávislé [19, 23].

Následně je transformační matice A převedena do její inverzní podoby W a je provedena lineární transformace původních signálů do vektoru maximálně nezávislých komponent podle vzorce 4.9 [23]:

$$s = Wx. \quad (4.9)$$

5 Klasifikátory

Klasifikátory rozhodují o tom, do jaké třídy budou analyzovaná data přiřazena. V našem případě se jedná o metody strojového učení. Na základě dostupných trénovacích dat jsou tyto systémy schopné učit se a tyto naučené znalosti pak využít při klasifikaci. Všechny níže popsané klasifikační metody se učí s učitelem, to znamená, že je zadán vstup a očekávaný výstup.

5.1 Podpůrné vektory

Metoda podpůrných vektorů je binární klasifikátor, tedy rozděljuje vstupní data do dvou tříd. Snaží se o nalezení nadroviny, která optimálně rozděljuje vstupní data v prostoru příznaků do dvou poloprostorů. Jednotlivé body reprezentují konkrétní příznaky. V případě nalezení optimálního řešení, body leží v poloprostorech a vzdálenost mezi těmito poloprostory, kde se nachází hledaná nadrovina, je co největší. Nadrovina je dána vzorcem 5.1:

$$y(x) = \omega^T x_k + \omega_0, \quad (5.1)$$

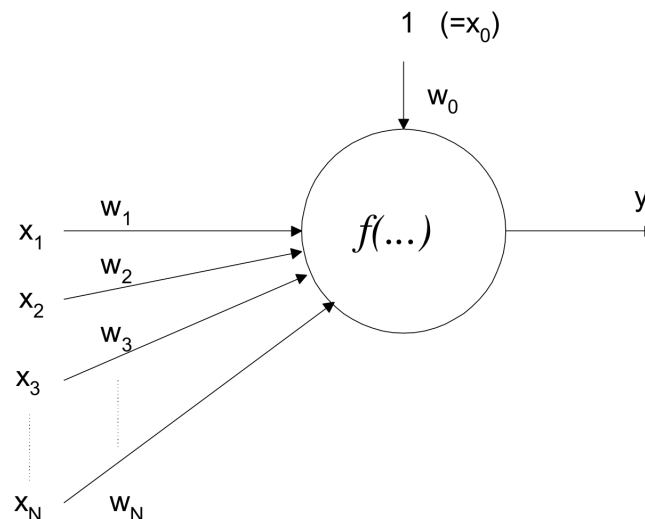
kde ω^T představuje orientaci přímky pomocí souřadnic vektoru ω , x_k je k -tý vstup. Poloha dané nadroviny je dána hodnotou prahu ω_0 . Nadrovina rozdělující příznaky do poloprostorů je lineární funkcí a je popsána nejbližšími body a body ležícími na její hranici, tyto body nazýváme podpůrné vektory [6].

5.2 Neuronové sítě

Neuronová síť je systém, který je složen z navzájem pospojovaných umělých neuronů (perceptronů) modelovaných jako biologické neurony se stejnou funkcí. Neuronová síť je tedy popsána jako zjednodušený model biologických neuronových sítí. Tento model je schopný správně zpracovávat informace a následně je transformovat do výstupních hodnot [17].

Správné nastavení parametrů je klíčové při vytváření modelu neuronové sítě. Obnáší vhodně zvolený počet vrstev sítě, počet neuronů v každé vrstvě, nastavené vzájemného propojení neuronů, dále výběr vhodné aktivační funkce, konstanty učení, způsobu učení a nastavení podmínek pro ukončení učení jako je například maximální počet epoch [17].

Na obrázku 5.1 je zobrazen umělý neuron definovaný vzorcem 5.2. x reprezentuje vstupy neuronu, w jsou váhy neuronu a w_0 prahovou hodnotu neboli bias. f je aktivační funkce, může být lineární, skoková, Heavisideova, sigmoidální nebo jako



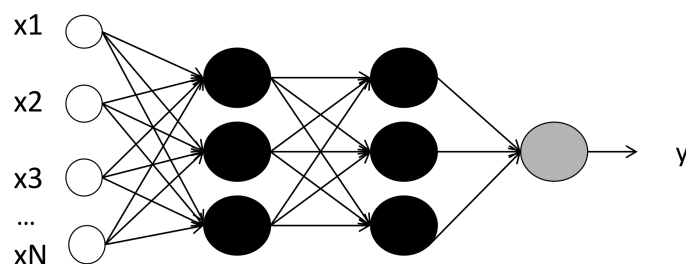
Obr. 5.1: Model perceptronu, převzato z: [17]

další například hyperbolická tangenta. Pomocí této funkce je transformován rozdíl váhovaného vstupu a prahu. y popisuje binární výstup neuronu [17],

$$y = f \left(\sum_{i=1}^N w_i x_i + w_0 \right). \quad (5.2)$$

Vícevrstvé perceptronové sítě

Vícevrstvé perceptronové sítě jsou dopředné neuronové sítě, které jsou nejvyžívanější. Ze vstupu se šíří informace pouze v dopředném směru až k výstupu a neurony nejsou mezi sebou rekurentně propojeny. Neuronová síť je tvořena nejméně třemi vrstvami: vstupní vrstva, skrytá vrstva a výstupní vrstva. Počet skrytých vrstev může být libovolný [17].



Obr. 5.2: Vícevrstvá perceptronová síť se dvěma skrytými vrstvami, převzato z:[16]

5.3 Lineární diskriminant

Lineární diskriminant je metoda založená na diskriminačních funkcích. Tyto funkce jsou odhadovány z trénovacích dat a jejich funkcí je zařazení objektu do optimální skupiny. Přiřazování do jednotlivých skupin probíhá pomocí vypočteného klasifikačního skóre. Výstupem diskriminační analýzy pak jsou skupiny s přiřazenými objekty [22].

5.4 K-nejbližší soused

Metoda k-nejbližších sousedů (k-NN, *k-nearest neighbour*) zařazuje objekty, které chceme klasifikovat, do dvou nebo více tříd. Zařazení probíhá na základě výpočtu euklidovské vzdálenosti D podle vztahu 5.3 mezi daným objektem x a každým prvkem příznakového vektoru \mathbf{v}_j trénovacích dat,

$$D = \sqrt{\sum_{i=1}^N (x(i) - \mathbf{v}_j(i))^2}. \quad (5.3)$$

Výpočet se provádí pro všechny příznakové vektory a výsledky jsou seřazeny od nejmenšího po největší. Objekt je následně zařazen do třídy, která se nejčastěji vyskytuje mezi jeho k-nejbližšími sousedy. Pokud bychom stanovili $k = 3$, hledáme tři nejbližší sousedy od objektu x a zařadíme ho do třídy, která převažuje mezi zvolenými sousedy [10].

6 Praktická část práce

6.1 Vstupní data

Tato práce pracuje s daty z databáze ÚBMI a je realizována v prostředí Mat1ab2018a. Jedná se o 10sekundové záznamy, které byly nasnímány 12svodovým EKG. Bylo naměřeno 8 svodů, konkrétně svody II, III a hrudní svody V1 - V6. Pro následnou analýzu byly vybrány signály ze svodu II. Součástí naměřených dat byly také detekované pozice QRS komplexů a anotace. Naměřená data jsou rozdělena do deseti skupin, fibrilace síní, sinusový rytmus, supraventrikulární tachykardie, komorová tachykardie, AV blokáda, blokáda Tawarových ramének, síňová extrasystola, komorová extrasystola, zašuměný signál a skupina ostatních. Dále byly zpracovány signály ze skupin síňová extrasystola a sinusový rytmus bez patologií.

6.1.1 Vybraná množina dat a její rozdělení

V dostupných datech se nachází 418 signálů se síňovou extrasystolou. Všechny tyto signály byly nakopírovány, aby byl získán větší počet signálů pro další zpracování. Dalším důvodem je rovnoměrné zastoupení signálů se sinusovým rytmem a síňovou extrasystolou. Nakopírováním bylo získáno 836 signálů se síňovou extrasystolou. Ze skupiny signálů se sinusovým rytmem bez patologií byl vybrán stejný počet signálů jako těch se síňovou extrasystolou, celkem tedy zpracováváme 1672 signálů.

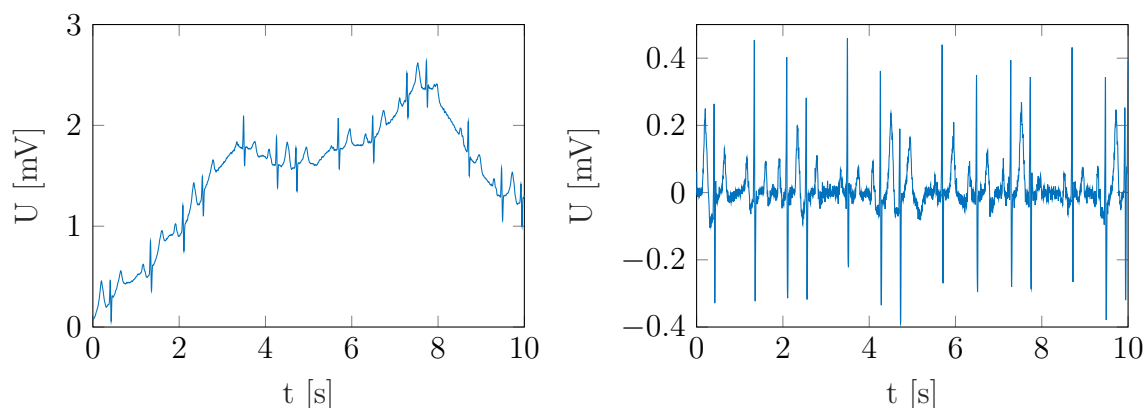
Rozdělení na trénovací a testovací množiny bylo provedeno 5násobnou křížovou validací. Data byla nejdříve rozdělena po pěti skupin, a poté byly signály se síňovou extrasystolou namnoženy, aby se nevyskytovaly v trénovací a testovací množině stejné signály. Ve všech skupinách se nachází přibližně 334 signálů se stejným počtem sinusových signálů a signálů se síňovou extrasystolou. Z těchto vybraných a rozdělených dat byla vytvořena nová množina dat, která je nadále zpracovávána.

6.1.2 Filtrace dat

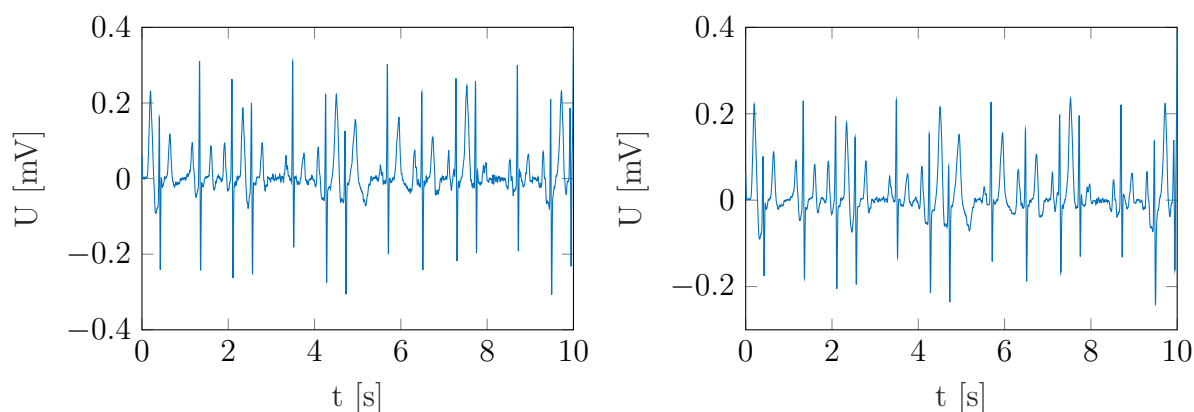
Data byla filtrována dvěma mediánovými filtry pro odstranění driftu. Nejprve byl použit mediánový filtr s délkou okna 200 ms, aby se odstranily QRS komplexy a P vlny. Poté se signál filtroval mediánovým filtrem o délce okna 600 ms, tím byly odstraněny T vlny a signál tak získal základní linii. Ta byla odečtena od původního signálu. Příklad signálu s driftem lze vidět na obrázku 6.1 vlevo.

Po odfiltrování driftu následovalo odstranění brumu. Filtrování proběhlo pomocí FIR filtru, konkrétně pásmovou zádrží s mezní frekvencí 50 Hz. Signál s brumem je na obrázku 6.1 vpravo.

Signál s odfiltrovaným driftem a brumem je znázorněn na obrázku 6.2 vlevo. Filtrace byla ještě doplněna o horní propust s mezní frekvencí 0,5-0,6 Hz z důvodu potřeby vyhlazení signálu pro následné zpracování. Výsledná filtrace signálu je zobrazena na obrázku 6.2 vpravo.



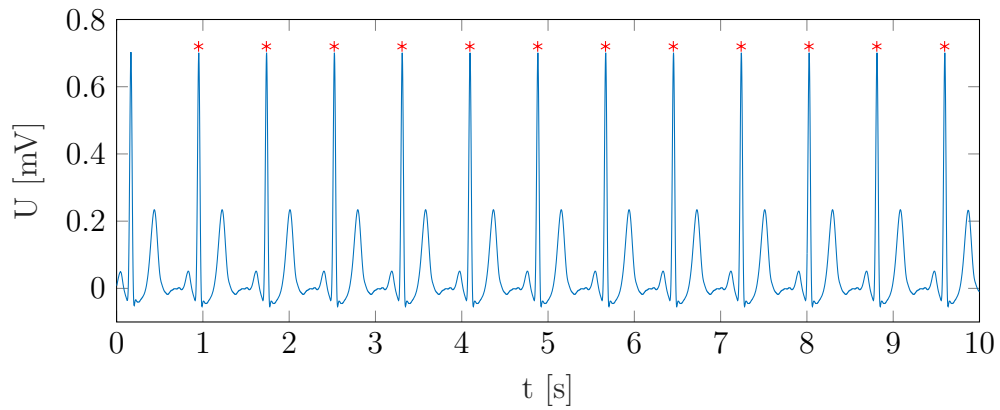
Obr. 6.1: Signál zatížený driftem (vlevo), signál zatížený brumem (vpravo)



Obr. 6.2: Signál s odfiltrovaným brumem a driftem, bez vyhlazení (vlevo), signál s filtrem na vyhlazení (vpravo)

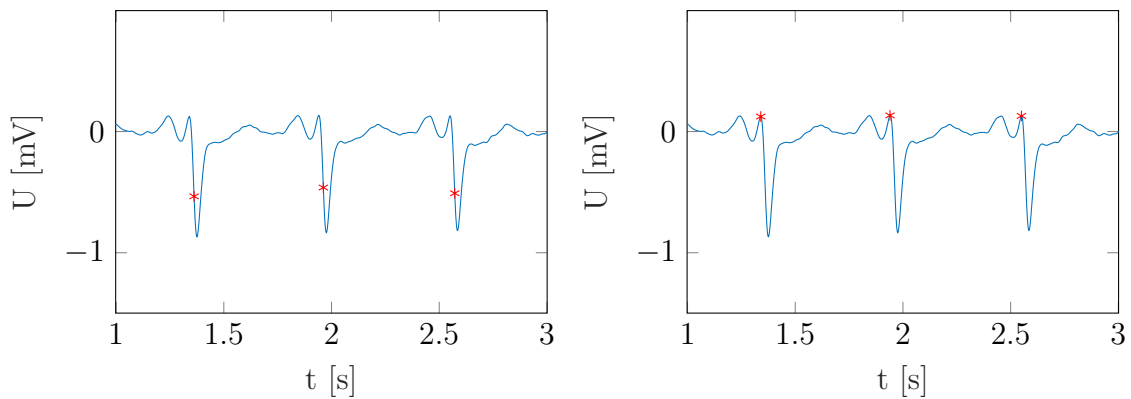
6.1.3 Detekované R kmity

Součástí analyzovaných dat jsou detekované R kmity, jejichž detekce v některých případech selhává a nejsou detekované všechny R kmity, které by měly být. Chybu detekce lze pozorovat na obrázku 6.3. Nedetekované R vlny byly vynechány, jednalo se především o krajní R vlny, nejčastěji první QRS komplexy. V rámci signálu nebyl daný EKG cyklus s chybějící detekcí R vlny dále zpracováván.



Obr. 6.3: Příklad chybné detekce R kmitů, první R kmit nebyl označen

Další případ chybné detekce je vidět na obrázku 6.4 vlevo, kde není detekován přesně R kmit, ale detekovaný bod se nachází v rámci QRS komplexu. Z tohoto důvodu bylo vytvořeno úzké okno o délce sinusového QRS komplexu a v tomto okně je detekovaný bod posunut na nejbližší maximální hodnotu, která je ve zvoleném okně. Správně nalezený R kmit je zobrazen na obrázku 6.4 vpravo.



Obr. 6.4: Špatně detekovaný R kmit (vlevo) a posunutá hodnota R kmitu na maximum (vpravo)

6.1.4 Segmentace dat

Nalezení bodů Q a S

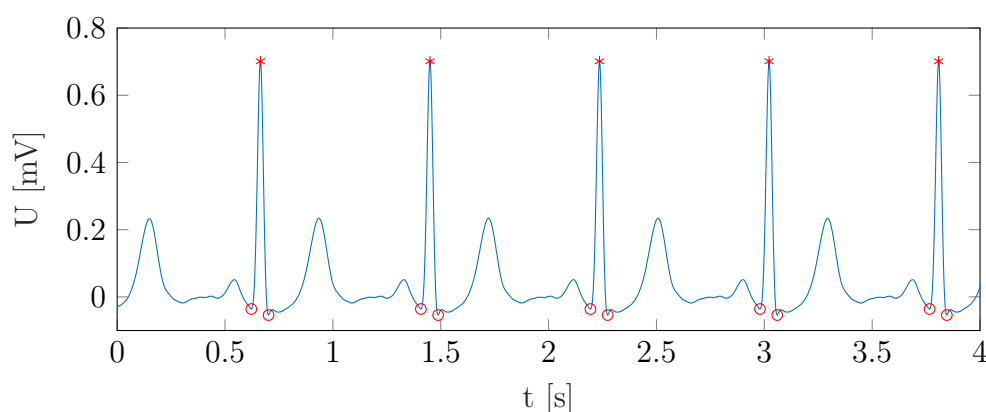
V rámci segmentace dat byly nejprve detekovány Q a S body. Pro nalezení těchto bodů byla zvolena dvě okna. Délka okna vychází z informace o délce normálního QRS komplexu, který trvá 0,04-0,11 s. Na základě zvolené vzorkovací frekvence byla vytvořena okna o délce 17 vzorků a 34 vzorků [27].

Pro každý bod Q byly nalezeny dvě hodnoty. Nejprve byl vyhledáván jako minimální hodnota v prvním okně, to je 17 vzorků od R vlny a byl označen jako bod Q1. Poté bylo okno zvětšeno na dvojnásobek prvního, tedy 34 vzorků od R vlny, a opět se bod Q2 hledal jako minimum. Následovalo nalezení maximální hodnoty v intervalu bodů Q1, Q2 a byla určena prahová hodnota [27].

Pokud se body Q1 a Q2 nacházejí na stejné pozici, jedná se o hledaný bod Q. V případě, že by bod Q2 měl větší amplitudu než bod Q1, pak je jako bod Q označeno Q1. Když je maximální hodnota z intervalu Q1, Q2 větší než součet amplitudy bodu Q1 a prahové hodnoty, poté je hledaným bodem Q2. Pokud nastane jiná možnost, je označen bod Q2 [27].

Stejným způsobem byl hledán bod S. Byla použita stejná délka oken, jako při hledání bodu Q. Ve dvou oknech o různých délkách byly určeny body S1 a S2. Byly nastaveny podmínky pro případ, že body S1 a S2 jsou stejné nebo, že je bod S1 je větší než S2. Při nesplnění ani jedné z uvedených podmínek, je hledaný bod S2 [27].

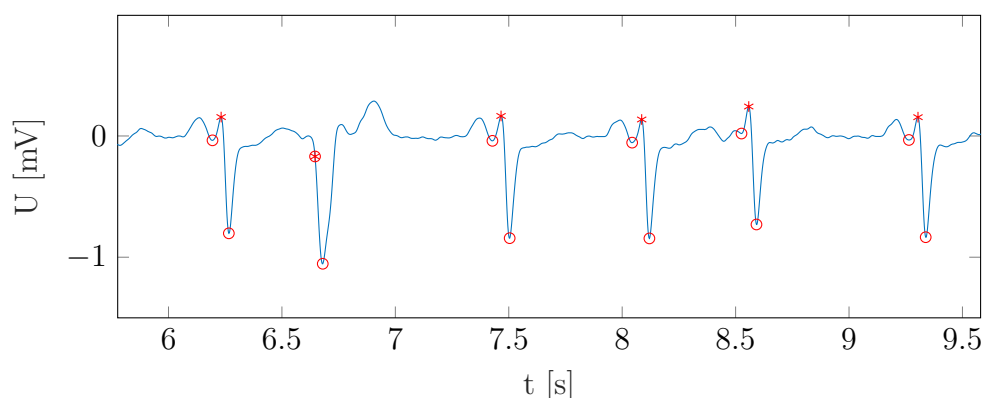
Pokud nastane u posledního cyklu případ, že je detekovaná R vlna a Q bod, ale bod S se zde nevyskytuje, tak je tento bod dopočítán. U předchozího EKG cyklu je vypočítána vzdálenost mezi bodem Q a S. Tato vzdálenost je připočítána k hodnotě Q bodu cyklu, u něhož dopočítáváme bod S. Detekované Q a S body jsou vidět na obrázku 6.5.



Obr. 6.5: Správně detekované Q a S body u sinusového EKG rytmu

Na EKG rytmu na obrázku 6.6 lze pozorovat jak správně nadetekované všechny body QRS komplexu u sinusových komplexů i komplexů se síňovou extrasystolou, tak i příklad špatné detekce. R vlny jsou označeny hvězdičkou a Q, S body kroužkem. V signálu se vyskytuje komorová extrasystola, kde je vidět špatně nadetekovaný R kmit, i když bylo aplikováno okno, kde bylo vyhledáváno maximum pro posunutí špatně detekovaného R kmitu na správnou hodnotu, R kmit stále není správně detekován. Byl sice posunut na hodnotu vyšší amplitudy, ale ne na maximální. Z tohoto

důvodu byla i špatně určená pozice Q bodu. V prvním i druhém okně pro nalezení tohoto bodu je minimální hodnota v těchto oknech stále pozice R vlny, proto je bod Q detekován na stejném místě jako R vlna. Řešením by bylo zvětšit velikost okna, ve kterém se maximum hledá, ale velikost okna je nastavena na šířku sinusového QRS komplexu, aby nejbližší maximum nebylo hledáno mimo QRS komplex. Toto úzké okno nestačí na vyhledání maxima v QRS komplexu komorové extrasystoly, která má oproti sinusovým signálům širší QRS komplex. V případě zvětšení okna by byla nejspíš R vlna u komorové extrasystoly detekována správně, ale u signálů s normálním QRS komplexem by mohlo docházet k chybným posunům, proto byla velikost okna ponechána.



Obr. 6.6: Chybná detekce Q a R bodu u komorové extrasystoly a správně nadetektované body QRS komplexu u sínové extrasystoly

Začátek a konec QRS komplexu

Po nalezení bodů Q a S byly určeny body označující začátek a konec QRS komplexu. Jedná se o odhad těchto bodů, protože byly nalezeny protnutím EKG křivky přímkou procházející nulou. Začátek QRS komplexu byl určen v prvním místě před Q bodem, kde křivka EKG prochází nulou a konec QRS komplexu jako první průchod nulou za bodem S. Příklady nalezení těchto bodů jsou na obrázcích 6.7 a 6.8

6.1.5 Výpočet příznaků

Z nalezených bodů byly vypočítány příznaky. Všechny tyto příznaky byly vypočítány pro každý průběh EKG v daném signálu. Následně bylo vybráno minimum, maximum, průměr a směrodatná odchylka. Vybrané hodnoty pro daný příznak poté reprezentovaly vždy jeden signál. Příznaky RR intervaly a průměrné lokální RR intervaly byly reprezentovány minimem, maximem, průměrem a směrodatnou odchylkou. Pro pozitivní a negativní plochu QRS komplexu a šířku QRS komplexu bylo

vybráno minimum, maximum a průměr. Přehledněji jsou vybrané příznaky vypsány v tabulce 6.1.

Tab. 6.1: Vybrané příznaky

RR intervaly (RR_int)	minimum maximum průměr směrodatná odchylka
Průměrné lokální RR intervaly (lokal_RR)	minimum maximum průměr směrodatná odchylka
Šířka QRS komplexu (šířka_QRS)	minimum maximum průměr
Pozitivní plocha QRS komplexu (auc_QRS+)	minimum maximum průměr
Negativní plocha QRS komplexu (auc_QRS-)	minimum maximum průměr

Z příznaků uvedených v tabulce 6.1 byl vytvořen příznakový vektor a data z něj byla standardizována. Standardizace byla provedena pomocí následujícího vzorce 6.1 pro výpočet Z-skóre,

$$z_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i}, \quad (6.1)$$

kde x_{ij} je j-tá hodnota i-tého příznaku, μ_i je střední hodnota daného příznaku a σ_i je směrodatná odchylka daného příznaku. Střední hodnoty i směrodatná odchylka konkrétního příznaku byly vypočítány zvlášť pro sinusové signály a pro signály se síňovou extrasystolou.

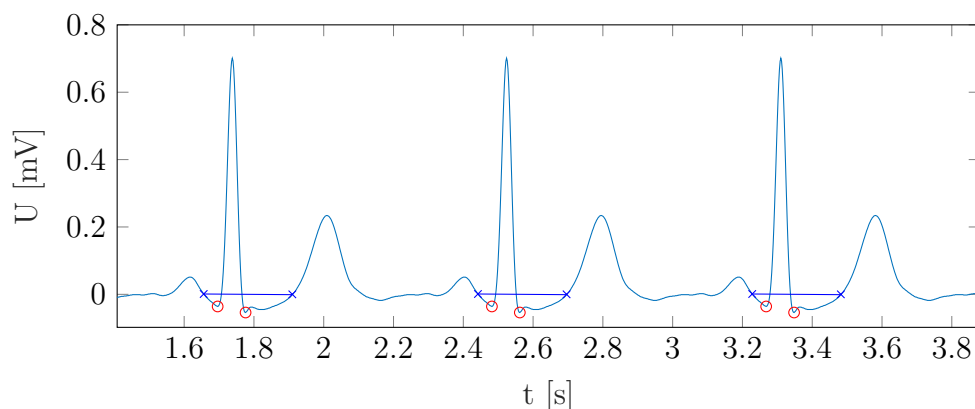
RR intervaly

Na základě nadetekovaných R vln byly vypočítány RR intervaly. RR interval byl vždy vypočítán mezi předchozí a aktuální R vlnou. Pro první hodnotu R vlny RR interval nelze dopočítat, proto tato hodnota byla nahrazena mediánem všech vypočtených RR intervalů z celého signálu.

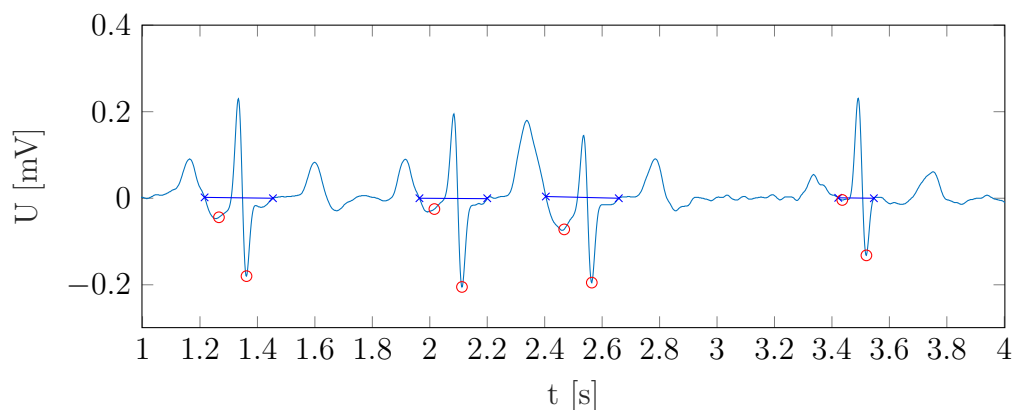
Poté byl vypočítán průměrný RR interval z celého záznamu a jako poslední lokální průměrný RR interval ze tří po sobě jdoucích RR intervalů, kdy jeho první a poslední hodnota byla nahrazena průměrem ze dvou aktuálních RR intervalů a průměrné hodnoty celého záznamu.

Šířka QRS komplexu

Šířka QRS komplexu byla vypočítána jako vzdálenost mezi začátkem a koncem odhadu QRS komplexu.



Obr. 6.7: Odhad začátku a konce QRS komplexu u sinusového EKG označené křížkem



Obr. 6.8: Označené začátky a konce QRS komplexu u signálu se sínovou extrasystolou.

Pozitivní a negativní plocha QRS komplexu

Pozitivní a negativní plocha QRS komplexu je rozdělená šířka QRS komplexu na kladnou a zápornou část, kdy jsou blíže popsány části průběhů sinusových i abnormálních QRS komplexů pro následující klasifikaci. Pokud bychom uvažovali sinusový QRS

komplex, skládal by se z první malé negativní vlny Q, která je následována velkou pozitivní vlnou R a poté malou negativní vlnou S. Výsledkem by bylo více pozitivní plochy než negativní. Signál se síňovou extrasystolou může mít i široký QRS komplex, z tohoto důvodu byla jeho pozitivní a negativní plocha zařazena mezi příznaky. Poté by šlo o velkou negativní vlnu Q a velkou pozitivní vlnu R. Ve velikosti pozitivní a negativní plochy by byl výrazně menší rozdíl než u sinusových signálů [24].

Plochy byly získány následujícím způsobem. Nejprve byly nalezeny body, kde křivka EKG prochází nulou, jednalo se o první průchod nulou před a za bodem Q. Stejně byly určeny tyto body před a za bodem S. Tyto čtyři body společně rozdělovaly QRS komplex na pozitivní a negativní část plochy.

7 Výsledky klasifikace metodou k-nejbližší soused

Jako klasifikační metoda pro detekci síňových extrasystol v EKG záznamech byla zvolena metoda k-nejbližšího souseda.

Po provedení 5násobné křížové validace byl příznakový vektor rozdělen do pěti částí, kdy 4 skupiny byly využity na učení klasifikátoru a zbývající skupina na testování. Celkem byl klasifikátor natrénován a otestován v pěti iteracích. Po každé iteraci byla vypočtena střední kvadratická chyba (MSE) a z těchto pěti hodnot byla získána průměrná hodnota střední kvadratické chyby z pěti provedených iterací.

Při učení klasifikátoru byl vstupem příznakový vektor složený ze čtyř skupin rozdělených křížovou validací a anotace k signálům, které jsou reprezentovány příznakovými vektory. Anotované signály s výskytem síňové extrasystoly jsou označeny jedničkou, sinusové signály bez patologií jsou označeny nulou, jedná se tedy o číselný vektor jedniček a nul.

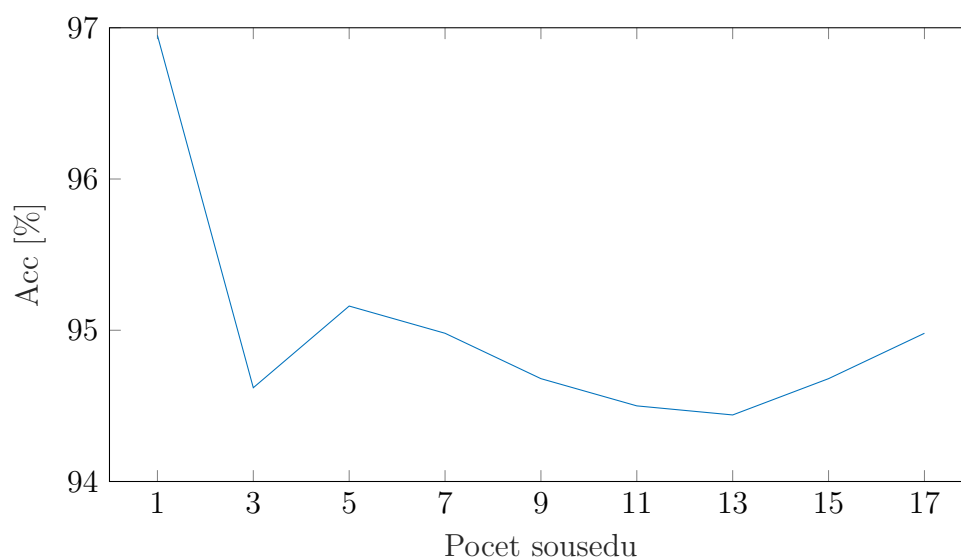
Dále byl při učení nastaven parametr počet nejbližších sousedů, na základě kterých se bude klasifikátor rozhodovat. Byla zvolena pouze lichá čísla, protože u sudých čísel může nastat případ, že ze dvou nejbližších sousedů je jeden signál s extrasystolou a druhý bez patologií. V případě lichých čísel k této situaci nemůže dojít. Jako maximální počet sousedů bylo zvoleno číslo 17, stejný počet je i zvolených příznaků, a zároveň odmocnina z počtu signálů na testování dosahuje také podobných hodnot [26]. Dále byla sledována chyba na výstupu klasifikátoru podle počtu nejbližších sousedů, která se s rostoucím počtem sousedů zvyšovala, stejně tak výpočetní náročnost algoritmu. Dalším parametrem bylo zvolení vzdálenostní metriky. Pro detekci byla vybrána euklidovská vzdálenostní metrika. Pro testování klasifikátoru byl vstup nastaven na jednu skupinu z pěti a její anotaci.

Po provedení trénování a testování byly vypočteny následující metriky: senzitivita (Se), specificita (Sp), pozitivní prediktivita (Pv+), přesnost (Acc), F1 skóre (F1), true positive (TP), true negative (TN), false positive (FP), false negative (FN) a střední kvadratická chyba klasifikátoru.

V tabulce 7.1 jsou uvedeny výsledky klasifikace pro příznakový vektor se všemi uvedenými příznaky. Je vidět, že nejlepší senzitivity, specificity, pozitivní prediktivity a přesnosti bylo dosaženo při klasifikaci pomocí pouze jednoho souseda, graf závislosti je zobrazen na obrázku 7.1. Zároveň bylo dosaženo nejmenší střední kvadratické chyby klasifikátoru s hodnotou 0,0305. Při použití různých počtů nejbližších sousedů nebylo v žádném případě dosaženo stejně vysokých nebo vyšších hodnot zvolených metrik.

Tab. 7.1: Výsledky klasifikace pro příznakový vektor s počtem sousedů pro lichá čísla od 1 do 17

	1	3	5	7	9	11	13	15	17
Se [%]	98,8	95,93	96,05	95,93	96,05	95,45	95,33	95,57	95,81
Sp [%]	95,1	93,3	94,26	94,02	93,3	93,54	93,45	93,78	94,14
Pv+ [%]	95,27	93,47	94,36	94,13	93,48	93,66	93,65	93,89	94,24
Acc [%]	96,95	94,62	95,16	94,98	94,68	94,5	94,44	94,68	94,98
F1	0,97	0,9468	0,9519	0,9502	0,9474	0,9454	0,9448	0,9472	0,9501
MSE	0,0305	0,0538	0,0484	0,0502	0,0532	0,055	0,0556	0,0532	0,0502



Obr. 7.1: Přesnost detekce v závislosti na počtu zvolených sousedů u klasifikátoru pro výpočet se všemi příznaky

7.1 Kombinace různých příznaků

Pro zvýšení úspěšnosti detekce byly zkoušeny různé kombinace příznaků, aby bylo dosaženo vyšších hodnot výsledků, a zároveň bylo zjištěno, jaké příznaky nejvíce ovlivňují správnost detekce, a které příznaky naopak úspěšnost detekce snižují. V tabulce 7.2 jsou uvedené všechny kombinace příznaků, pro které byla prováděna detekce, jsou buď s jedním chybějícím příznakem nebo se dvěma.

Tab. 7.2: Vybrané kombinace příznaků

Varianty	Vybrané kombinace příznaků
1	lokal_RR, šířka_QRS, auc_QRS+, auc_QRS-
2	RR_int, šířka_QRS, auc_QRS+, auc_QRS-
3	RR_int, lokal_RR, auc_QRS+, auc_QRS-
4	RR_int, lokal_RR, šířka_QRS, auc_QRS-
5	RR_int, lokal_RR, šířka_QRS, auc_QRS+
6	RR_int, auc_QRS+, auc_QRS-
7	RR_int, lokal_RR, auc_QRS+
8	RR_int, lokal_RR, auc_QRS-

Tabulka 7.3 uvádí vypočtené hodnoty senzitivity, specifity, pozitivní prediktivity, přesnosti a středné kvadratické chyby. Tyto hodnoty byly vypočítány pro všechny uvedené příznaky z tabulky 7.2. U dané varianty je vždy uveden výsledek klasifikace s jedním sousedem, protože i případě vybraných příznaků dosáhla metoda k -nejbližšího souseda nejlepších výsledků detekce s hodnotou $k = 1$.

Tab. 7.3: Výsledky klasifikace pro různé varianty nakombinovaných příznaků s jedním sousedem

	1	2	3	4	5	6	7	8
Se [%]	96,17	98,56	99,52	98,56	98,32	99,28	98,56	99,28
Sp [%]	88,03	95,33	96,17	95,81	95,69	96,05	96,53	97,12
Pv+ [%]	88,93	95,48	96,29	95,92	95,8	96,17	96,6	97,18
Acc [%]	92,1	96,94	97,84	97,18	97	97,66	97,54	98,2
F1	0,9241	0,9699	0,9788	0,9722	0,9704	0,977	0,9757	0,9822
MSE	0,0789	0,0305	0,0215	0,0281	0,0299	0,0233	0,0245	0,0179

Nejlepších výsledků bylo dosaženo 8. variantou kombinací příznaků, jednalo se o příznakový vektor složený z RR intervalů, průměrných lokálních RR intervalů

a negativních ploch QRS komplexů signálu. Výsledky detekce jsou uvedeny v tabulce 7.3. Vzhledem k vysokým hodnotám metrik reprezentující výsledky klasifikace lze usuzovat, že právě vynechané příznaky jako šířka QRS komplexu a pozitivní plocha QRS komplexu se ve velké míře nepodílí na správné detekci signálů, protože ve srovnání s výsledky s kompletním příznakovým vektorem je MSE podstatně nižší a přesnost detekce vyšší.

Když se podíváme blíže na příznaky, které byly vynechány, šířka QRS komplexu (3. varianta) je příznakem, který výsledky detekce zhoršuje v porovnání s původním příznakovým vektorem se všemi příznaky. Šířka QRS komplexu dostatečně nepopisuje data natolik, aby příznak přispíval ke správné detekci, naopak detekci zhoršuje a mate klasifikátor. Nejspíš je to způsobeno výpočtem tohoto příznaku, protože se jedná pouze o odhad šířky QRS komplexu.

Dalším příznakem je pozitivní plocha QRS komplexu (4. varianta). Na základě uvedených výsledků lze pozorovat, že bez tohoto příznaku výsledky klasifikace dosahují podobných výsledků jako se všemi příznaky. Z toho vyplývá, že se jedná o příznak, který by mohl být vynechán, protože klasifikátoru neposkytuje relevantní informace o datech, na základě kterých by mohlo dojít k výslednému rozhodnutí o výskytu síňových extrasystol.

Klasifikace pomocí příznakového vektoru bez přítomnosti hodnot o šířce QRS komplexů a lokálních RR intervalů také dosahovala vysokých detekčních výsledků, stejně tak kombinace příznaků bez šířky QRS komplexu a negativní plochy QRS komplexu. Vzhledem k tomu, že součástí těchto kombinací je i šířka QRS komplexu, který výsledky klasifikace ztlačuje, jedná se o důvod, proč příznakový vektor bez těchto příznaků dosahuje lepších výsledků, než s nimi. Pokud z příznakového vektoru vynecháme průměrný lokální interval nebo negativní plochu QRS komplexu, výsledky detekce se pohybují kolem stejných hodnot, jako když jsou tyto příznaky přítomny.

Nejvíce se na správné detekci síňových extrasystol podílí výpočet RR intervalů. V případě, že z příznakového vektoru tento důležitý příznak odstraníme, dochází ke ztlačení klasifikace, kdy přesnost klesá o necelých pět procent a střední kvadratická chyba je zvýšena z hodnoty 0,0305 na hodnotu 0,0789 oproti kompletnímu příznakovému vektoru, kterého je tento příznak součástí. Významnou vlastností síňových extrasystol je neúplná kompenzační pauza, kdy jsou RR intervaly nepravidelné oproti sinusovým signálům, které mají RR intervaly pravidelné, z tohoto důvodu je tento příznak klíčový pro správnou klasifikaci.

Hodnoty reprezentující výsledky testování naučeného klasifikátoru v tabulkách 7.1 a 7.3 sice vypovídají o tom, jakou mírou se který příznak podílí na správné detekci síňových extrasystol, a který lépe reprezentuje vlastnosti jednotlivých signálů, ale zároveň záleží na tom, jak byl daný klasifikátor naučen. U tabulky 7.3 jde s jisto-

tou říci, že příznaky odvozené z RR intervalů se na správné detekci podílí nejvíce ze všech uvedených příznaků. U varianty číslo 8 lze usuzovat, že vynecháním příznaků šířka QRS komplexu a pozitivní plocha QRS komplexu dosáhneme vyšších výsledků než při použití kompletního příznakového vektoru. Když porovnáme střední kvadratickou chybu mezi zbylými variantami kombinací, je vidět, že hodnoty si jsou velmi blízké. Z těchto kombinací proto nejde říci, která kombinace je lepší a bude dosahovat vyšších výsledků, protože výsledky jsou ovlivněny tím, jak se klasifikátor naučí a na základě toho se rozhodne.

7.2 Metoda PCA

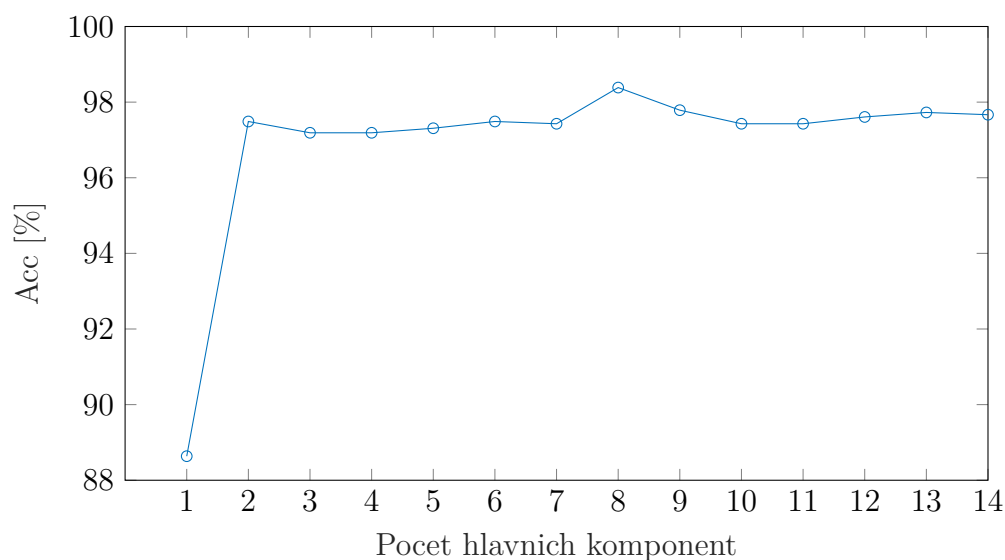
Na vytvořené příznakové vektory byla aplikována metoda PCA. Cílem bylo transformovat původní příznakové vektory do prostoru hlavních komponent a získat vzájemně nekorelované příznaky. Klasifikace byla opět provedena pro jednoho souseda. V tabulce 7.4 jsou uvedeny výsledky klasifikace po provedení PCA.

Tab. 7.4: Výsledky klasifikace s jedním sousedem po aplikaci PCA pro varianty příznaků

	1-8	2	3	4	5	6	7	8
Se [%]	99,52	99,28	99,52	100	99,76	99,28	99,28	99,04
Sp [%]	96,77	95,45	97,12	96,77	96,77	96,05	96,77	97,12
Pv+ [%]	96,85	95,62	97,19	96,87	96,86	96,17	96,84	97,18
Acc [%]	98,14	97,36	98,32	98,38	98,26	97,66	98,02	98,08
F1	0,9817	0,9741	0,9834	0,9841	0,9829	0,9771	0,9805	0,9811
MSE	0,0185	0,0263	0,0167	0,0161	0,0173	0,0233	0,0197	0,0191

Pro kombinaci příznaků bez přítomnosti RR intervalů nebylo aplikováno PCA, protože detekce bez této informace dosahuje vysoké chybovosti, proto tato kombinace příznaků není dále uvažována jako vhodná pro klasifikaci síňových extrasystol.

Na základě výsledků je vidět, že pro 6. kombinaci příznaků jsou výsledky detekce stejné jako u klasifikace bez provedení PCA, proto není vhodné počítat hlavní komponenty pro tato data. U 8. skupiny příznaků bylo dosaženo horších výsledků v porovnání s detekcí bez PCA. Na tyto příznaky by se dalo nahlížet jako na jejich vhodnou kombinaci, která by obsahovala vzájemně nezávislé příznakové vektory. Na druhou stranu se jedná o případ, kdy analýza hlavních komponent selhala a není doporučena.



Obr. 7.2: Přesnost klasifikátoru v závislosti na počtu hlavních komponent pro 4. kombinaci příznaků

Pokud budeme uvažovat nejlepší výsledky, jednalo by se o 4. variantu s přesností 98,38 %. Této hodnoty bylo dosaženo osmi hlavními komponentami, jak je vyznačeno na obrázku 7.2, které reprezentují data z 99,5 %.

Pro ostatní uvedené možnosti příznaků dochází ke zlepšení výsledků klasifikace díky odstranění příznaků reprezentující stejnou informaci v datech. V těchto konkrétních kombinacích se především jedná o podobnost mezi šířkou QRS komplexu, pozitivní a negativní plochou QRS komplexu. Jistá podobnost se vyskytuje i mezi RR intervaly a průměrnými lokálními RR intervaly.

8 Diskuze

Detekce signálů s přítomností síňových extrasystol byla provedena pomocí klasifikátoru k-nejbližší sused, který je vhodný pro menší dataset. Nejlepších výsledků bylo dosaženo pouze s jedním susedem a euklidovskou vzdálenostní metrikou. Klasifikátor se rozhodoval o zařazení do třídy podle jednoho nejbližšího suseda, důvodem mohla být podobnost mezi analyzovanými daty, která vznikla namnožením signálů se síňovou extrasystolou z důvodu nízkého počtu těchto signálů.

Následovalo nalezení vhodné kombinace příznaků, která by dosahovala co nejvyšších výsledků detekce. Na všechny tyto příznakové vektory byla aplikována metoda PCA pro snížení závislosti mezi zvolenými příznaky, a tím i zlepšit výslednou detekci.

Tab. 8.1: Výsledky klasifikace pomocí k-NN a s PCA pro všechny příznaky v horní části tabulky a pro 5. kombinaci příznaků v dolní části tabulky

	Se	Sp	Pv+	Acc
k-NN	98,8	95,1	95,27	96,95
k-NN + PCA	99,5	96,77	96,85	98,14
k-NN	98,32	95,69	95,8	97
k-NN + PCA	99,76	96,77	96,86	98,26

V případě použití PCA u všech příznaků i 5. varianty je zvýšen počet správně detekovaných signálů, a zároveň se snižuje počet falešně negativních a pozitivních výsledků. Díky metodě PCA bylo dosaženo úspěšnější detekce, jak lze vidět na hodnotách přesnosti klasifikátoru v tabulce 8.1. Klasifikace se všemi vypočítanými příznaky metodou k-NN a při výpočtu PCA ukázala na redundantnost některých příznaků a jejich vzájemnou korelaci, což jsou vlastnosti, které na správnou detekci působí negativně. Pátá varianta je bez výpočtu negativní plochy QRS komplexu. Příznak šířka QRS komplexu je vypočítána jako odhad a pomocí této hodnoty je dopočítána negativní a pozitivní plocha QRS komplexu, to může být důvodem jejich nedůležitosti při popisu dat. Tyto příznaky mezi sebou vykazují podobnost, která je metodou PCA redukována a výsledkem je úspěšnější detekce.

Pro 4. variantu vybraných příznaků bylo dosaženo nejlepších detekčních výsledků navrženého algoritmu, výsledky jsou uvedeny v tabulce 8.2. PCA se podílí na zlepšení výsledné klasifikace. Pro čtvrtou kombinaci příznaků po aplikaci výpočtu hlavních komponent bylo dosaženo nulového počtu falešně negativních výsledků a senzitivity 100 %. Tato kombinace je bez přítomnosti příznaku pozitivní plocha QRS komplexu.

Pokud bychom analyzovaná data chtěli klasifikovat pomocí výše diskutovaných kombinací příznaků, bylo by vhodné zahrnout analýzu hlavních komponent. Podle

Tab. 8.2: Nejlepší dosažené výsledky klasifikace pomocí k-NN a s PCA pro 4. kombinaci příznaků

	Se	Sp	Pv+	Acc
k-NN	98,56	95,81	95,92	97,18
k-NN + PCA	100	96,77	96,87	98,38

výsledků klasifikace pro dané příznaky lze vidět, že mezi příznaky, které nedostatečně popisují data patří negativní a pozitivní plocha QRS komplexu. Také je známa jejich závislost na šířce QRS komplexu. Při použití PCA je dosaženo vyšších detekčních výsledků, protože metoda tyto příznaky transformuje na nové nekorelované.

Tab. 8.3: Výsledky klasifikace pomocí k-NN a s PCA pro 3. kombinaci v horní části tabulky a pro 7. kombinaci v dolní části tabulky

	Se	Sp	Pv+	Acc
k-NN	99,52	96,17	96,29	97,84
k-NN + PCA	99,52	97,12	97,19	98,32
k-NN	98,56	96,53	96,6	97,54
k-NN + PCA	99,28	96,77	96,84	98,02

U kombinací příznaků s vynecháním šířky QRS komplexu nebo společně s negativní plochou QRS komplexu dochází k minimálním změnám ve výsledcích detekce, když porovnáme výsledky dosažené k-NN a s PCA. Výsledky jsou uvedeny také v tabulce 8.3. Důvodem je odstranění příznaků, které se ve větší míře nepodílí na výsledné detekci. Tyto příznaky částečně popisují stejné informace v datech, proto bez jejich přítomnosti je zároveň snížena podobnost mezi jednotlivými příznaky a metoda PCA dosahuje podobných výsledků jako bez jejího použití.

Detekcí pomocí 6. varianty příznaků bylo dosaženo stejných výsledků bez rozdílu v tom, zda bylo použito PCA. Matice příznaků byla složena z RR intervalů, pozitivní a negativní plochy QRS komplexu. V tomto případě je zbytečné používat PCA, protože s počítáním hlavních komponent se zvyšuje výpočetní náročnost algoritmu a výsledky zůstávají stejné. To platí také u 2. i 8. kombinace příznaků, jak lze usuzovat podle výsledků v tabulkách 7.3 a 7.4. V případě těchto kombinací příznaků jsou dostatečně popsána data, aby mohla být správně klasifikována bez použití redukčních metod. Analýza hlavních komponent je pro tyto příznaky považována za redundantní a je vhodné použít klasifikátor bez PCA.

Závěr

Cílem práce bylo navrhnout a implementovat systém k detekci síňových extrasystol v EKG signálech.

Na začátku práce je popsána činnost srdce, EKG křivka a fyziologie síňových extrasystol vyskytujících se v těchto záznamech. Druhá kapitola pojednává o typech rušení, které znehodnocují signál, který chceme zpracovávat. Jsou zde popsány i postupy, kterými jsou rušení odstraňována na základě dostupných studií. Další část obsahuje informace o segmentaci signálu. Ve třetí kapitole je psáno o výběru příznaků, jakými způsoby je získáváme a jak příznaky rozdělujeme. Čtvrtá kapitola je zabývá metodami na redukci příznaků. Součástí paté kapitoly je shrnutí klasifikačních metod pro detekci síňových extrasystol.

Následující kapitola se zabývá praktickou částí bakalářské práce, kde je popsán postup, jakým byl navržen algoritmus pro detekci síňových extrasystol. Popisuje vstupní data, jejich předzpracování a segmentaci. Součástí je tvorba příznakového vektoru a jeho redukce pomocí analýzy hlavních komponent. Dále popsání jednotlivých skupin, do kterých byly příznakové vektory rozděleny. Práce je zaměřena na klasifikátor k-nejbližšího souseda, který byl implementován v prostředí `Matlab`.

Poslední kapitola obsahuje výsledky klasifikátoru k-NN pro všechny analyzované skupiny příznaků a redukované příznaky analýzou hlavních komponent. Nejlepší výsledky v detekci síňových extrasystol vykazoval klasifikátor s přesností 98,38 %. Těchto výsledků bylo dosaženo s nastavením počtu sousedů na hodnotu 1 a s euklidovskou vzdáleností. Se zvyšujícím se počtem sousedů se přesnost detekce snižovala. Mezi vybrané příznaky byly zařazeny RR intervaly, lokální průměrné RR intervaly, šířka QRS komplexu a také jeho negativní plocha. Tato skupina příznaků se ukázala jako nejefektivnější pro řešení daného problému společně s výpočtem hlavních komponent.

Nejhůře si vedl klasifikátor v případě, že příznakový vektor neobsahoval informace o RR intervalech. V tomto případě se úspěšnost pohybovala na 92 %. Z výsledků je zřejmá důležitost tohoto příznaku pro detekci síňových extrasystol.

Důležitým faktorem ovlivňující výsledky detekce je použitý dataset, který obsahoval nižší zastoupení signálů s přítomností síňových extrasystol, což se pozitivně projevilo na výpočetní náročnosti algoritmu, ale zároveň klasifikátor nemusel mít dostatek informací pro přesnější zařazení. Důležitým krokem je i správný výběr příznaků reprezentující dané signály. Jako zavádějící se především ukázaly příznaky jako šířka QRS komplexu společně s jeho pozitivní a negativní plochou v případě, že není počítáno PCA. Důvodem nepřesnosti popisu dat u těchto příznaků je především fakt, že se jedná o přibližný odhad jejich hodnot.

Výsledky této práce by bylo možné vylepšit použitím obsáhlejšího datasetu nebo

přesnějším výpočtem příznaků.

Literatura

- [1] SILBERNAGL, Stefan a Agamemnon DESPOPOULOS. *Atlas fyziologie člověka: překlad 8. německého vydání*. 4. české vydání. Praha: Grada Publishing, 2016. ISBN 978-80-247-4271-7.
- [2] BURNS, Ed. *Premature Atrial Complex (PAC)* [online]. Jun 5, 2019 [cit. 2020-10-22]. <https://litfl.com/premature-atrial-complex-pac/>
- [3] CEYLAN, Rahime a Yüksel ÖZBAY. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. *Expert systems with applications* [online]. Elsevier, 2007, **33**(2), 286-295 [cit. 2020-11-28]. ISSN 0957-4174. doi:10.1016/j.eswa.2006.05.014
- [4] ČIHALÍK, Čestmír a Miloš TÁBORSKÝ. *EKG v klinické praxi*. Olomouc: Solen, Medical education, 2013. Meduca. ISBN 978-80-7471-015-5.
- [5] DE LANNOY, Gaël, Damien FRANCOIS, Jean DELBEKE a Michel VERLEYSEN. Weighted Conditional Random Fields for Supervised Interpatient Heartbeat Classification. *IEEE transactions on biomedical engineering* [online]. IEEE, 2012, **59**(1), 241-247 [cit. 2020-11-07]. ISSN 0018-9294. doi:10.1109/TBME.2011.2171037
- [6] DE LANNOY, G., D. FRANÇOIS, J. DELBEKE a M. VERLEYSEN. Weighted SVMs and feature relevance assessment in supervised heart beat classification. In: *Communications in Computer and Information Science* [online]. 2011, s. 212-223 [cit. 2020-12-18]. ISBN 9783642184710. ISSN 18650929. doi:10.1007/978-3-642-18472-7_17
- [7] GOLDBERGER, Ary Louis, Zachary D. GOLDBERGER a Alexei SHVILKIN. *Goldberger's clinical electrocardiography: a simplified approach*. Ninth edition. Philadelphia, PA: Elsevier, [2018]. ISBN isbn:978-0-323-40169-2.
- [8] HABERL, Ralph. *EKG do kapsy*. Praha: Grada, 2012. ISBN 978-80-247-4192-5.
- [9] KOZUMPLÍK, Jiří. *Analýza biologických signálů 1*. Brno: VUT FEKT, 2018. Skripta.
- [10] LANATÁ, A, G VALENZA, C MANCUSO a E.P SCILINGO. Robust multiple cardiac arrhythmia detection through bispectrum analysis. *Expert systems with applications* [online]. Elsevier, 2011, **38**(6), 6798-6804 [cit. 2020-12-18]. ISSN 0957-4174. doi:10.1016/j.eswa.2010.12.066

- [11] LIN, Chun-cheng a Chun-min YANG. Heartbeat Classification Using Normalized RR Intervals and Morphological Features. *Mathematical Problems in Engineering* [online]. Hindawi Publishing Corporation, 2014, **2014** [cit. 2020-11-07]. ISSN 1024-123X. doi:10.1155/2014/712474
- [12] LLAMEDO SORIA, M. a J.P. MARTÍNEZ. Analysis of multidomain features for ECG classification. In: *Computers in Cardiology* [online]. 2009, s. 561-564 [cit. 2020-11-07]. ISBN 9781424472819. ISSN 02766574.
- [13] LUZ, Eduardo José Da S, William Robson SCHWARTZ, Guillermo CÁMARA-CHÁVEZ a David MENOTTI. ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine* [online]. Elsevier Ireland, 2016, **127**(C), 144-164 [cit. 2020-11-06]. ISSN 0169-2607. doi:10.1016/j.cmpb.2015.12.008
- [14] O'DWYER, M a R.B REILLY. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE transactions on biomedical engineering* [online]. IEEE, 2004, **51**(7), 1196-1206 [cit. 2020-11-07]. ISSN 0018-9294. doi:10.1109/TBME.2004.827359
- [15] RAWSHANI, Araz. *Clinical ECG Interpretation* [online]. [cit. 2020-11-06]. <https://ecgwaves.com/topic/premature-atrial-contraction-beat-complex/>
- [16] RONZHINA, Marina. *BPC-UIM - Umělá inteligence v medicíně: Umělé neuronové sítě*. Brno, 2020. Přednáška. VUT FEKT.
- [17] KOZUMPLÍK, Jiří a Ivo PROVAZNÍK. *Umělá inteligence v medicíně*. Brno: VUT FEKT, 2007.
- [18] HARUŠTIAKOVÁ, Danko. *Vícerozměrné statistické metody v biologii*. Brno: Akademické nakladatelství CERM, 2012. ISBN 978-80-7204-791-8.
- [19] SARFRAZ, Mohammad, Ateeq Ahmed KHAN a Francis F. LI. Using independent component analysis to obtain feature space for reliable ECG Arrhythmia classification. In: qtextit2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) [online]. IEEE, 2014, 2014, s. 62-67 [cit. 2020-12-03]. ISBN 978-1-4799-5669-2. doi:10.1109/BIBM.2014.6999249
- [20] THALER, Malcolm S. *EKG a jeho klinické využití*. Praha: Grada, 2013. ISBN 978-80-247-4193-2.
- [21] WILLACY, Hayley. *Extrasystoles* [online]. 26. 5. 2019 [cit. 2020-10-03]. <https://patient.info/doctor/extrasystoles#>

- [22] FRIEDMAN, Jerome, Trevor HASTIE a Robert TIBSHIRANI, *The Elements Of Statistical Learning* (2001), Springer Series in Statistics.
- [23] YU, S a K CHOU. Integration of independent component analysis and neural networks for ECG beat classification. *Expert Systems with Applications* [online]. 2008, **34**(4), 2841-2846 [cit. 2020-12-03]. ISSN 09574174. doi:10.1016/j.eswa.2007.05.006
- [24] RAWSHANI, Araz. *The QRS complex: ECG features of the Q-wave, R-wave, S-wave & duration* [online]. [cit. 2021-05-15]. <https://ecgwaves.com/ecg-qrs-complex-q-r-s-wave-duration-interval/>
- [25] STERNICKEL, Karsten. Automatic pattern recognition in ECG time series. *Computer Methods and Programs in Biomedicine* [online]. 2002, **68**(2), 109-115 [cit. 2020-12-3]. ISSN 01692607. doi:10.1016/S0169-2607(01)00168-7
- [26] ISHANI, Patel, Virag JAGTAP a Ompriya KALE. Article: A Survey on Feature Extraction Methods for Handwritten Digits Recognition. *International Journal of Computer Applications*. 2014, **107**(12):11-17
- [27] YEH, Yun-chi a Wen-june WANG. QRS complexes detection for ECG signal: The Difference Operation Method. *Computer methods and programs in bio-medicine* [online]. Elsevier Ireland, 2008, **91**(3), 245-254 [cit. 2021-5-18]. ISSN 0169-2607. doi:10.1016/j.cmpb.2008.04.006

Seznam symbolů, veličin a zkratek

SA	sinoatriální
AV	atrioventrikulární
EKG	elektrokardiogram
FIR	konečná impulzní odezva
DWT	diskrétní vlnková transformace
PCA	analýza hlavních komponent
ICA	analýza nezávislých komponent
k-NN	k-nejbližší soused
Se	senzitivita
Sp	specifita
Pv+	pozitivní prediktivita
Acc	přesnost
F1	F1 skóre
TP	true positive
TN	true negative
FP	false positive
FN	false negative
MSE	střední kvadratická chyba
RR_int	RR intervaly
lokal_RR	průměrné lokální RR intervaly
šířka_QRS	šířka QRS komplexu
auc_QRS+	pozitivní plocha QRS komplexu
auc_QRS-	negativní plocha QRS komplexu